**DEPARTMENT OF ECONOMICS
DISCUSSION PAPER SERIES**

**FACTOR HIGH-FREQUENCY BASED VOLATILITY (HEAVY)
MODELS**

Kevin Sheppard and Wen Xu

**Number 710
May 2014**

Manor Road Building, Manor Road, Oxford OX1 3UQ

Factor High-Frequency Based Volatility (HEAVY) Models

Kevin Sheppard*

Department of Economics & Oxford-Man Institute
University of Oxford
kevin.sheppard@economics.ox.ac.uk

Wen Xu

Department of Economics & Oxford-Man Institute
University of Oxford
wen.xu@economics.ox.ac.uk

May 27, 2014

Abstract

We propose a new class of multivariate volatility models utilizing realized measures of asset volatility and covolatility extracted from high-frequency data. Dimension reduction for estimation of large covariance matrices is achieved by imposing a factor structure with time-varying conditional factor loadings. Statistical properties of the model, including conditions that ensure covariance stationary or returns, are established. The model is applied to modeling the conditional covariance data of large U.S. financial institutions during the financial crisis, where empirical results show that the new model has both superior in- and out-of-sample properties. We show that the superior performance applies to a wide range of quantities of interest, including volatilities, covolatilities, betas and scenario-based risk measures, where the model's performance is particularly strong at short forecast horizons.

Keywords: Conditional Beta, Conditional Covariance, Forecasting, HEAVY, Marginal Expected Shortfall, Realized Covariance, Realized Kernel, Systematic Risk

JEL Classification: C32, C53, C58, G17, G21

*We thank seminar and conference participants at Bath University, the Office of Financial Research, EUI, and SoFiE 2013 (Singapore) for helpful comments. Code for computing the realized measures used in this paper is available at <http://www.kevinsheppard.com>.

1 Introduction

Conditional covariances are key inputs in risk management and portfolio optimization. Traditionally, multivariate GARCH models based on daily data are used to capture the dynamics of the second-order moments of asset returns (see [Bauwens, Laurent & Rombouts 2006](#) for a survey), and while most multivariate volatility models are feasible when the number of assets is small – 5 or fewer – only a small subset remain feasible when applied to large, empirically realistic portfolios. There are a number of difficulties in high-dimension covariance modeling, including the computational effort required to invert large conditional covariance matrices when evaluating the likelihood and the high-dimensionality of the parameter space. Recent contributions to the literature have attempted to side-step these issues by using alternative estimators or carefully designed models. [Engle, Shephard & Sheppard \(2008\)](#) construct the composite likelihood by summing up the log-likelihoods of pairs of assets in order to avoid the inversion of high-dimensional matrices. The Dynamic Equicorrelation model proposed in [Engle & Kelly \(2012\)](#) leads to a simple analytic form of the inverse of the conditional covariance by assuming that the time-varying correlations are identical across all pairs of assets. A third, and older, approach to achieve dimension reduction is to exploit the strong factor structure of asset returns when modeling conditional covariance. Early examples include [Engle, Ng & Rothschild \(1990\)](#), who introduced the factor-ARCH model to price Treasury bills (see also [Diebold & Nerlove \(1989\)](#) and [Ng, Engle & Rothschild \(1992\)](#)).

Recently, intra-daily estimators of volatility – collectively known as realized measures – have been used to improve volatility models. Realized measures incorporate information from asset price paths to improve the measurement of volatility over a fixed horizon, typically one day. The simplest and most common realized measure is realized variance, which estimates the quadratic variation of the intra-daily log-price process using the sum of the squared high-frequency returns. When the prices follows a diffusion with stochastic volatility and can be directly observed without error, realized variance converges to daily integrated variance of the underlying volatility process ([Andersen, Bollerslev, Diebold & Labys 2001](#); [Barndorff-Nielsen & Shephard 2002](#)). However, microstructure noise such as bid-ask bounce is ubiquitous in high frequency data which limits sampling frequency. Several alternative solutions exist in the literature to control microstructure effect, including two- and multi-scale realized volatility ([Aït-Sahalia et al. 2005](#), [Zhang 2006](#)), realized kernel estimators ([Barndorff-Nielsen, Hansen, Lunde & Shephard 2008](#)), and the pre-averaging approach ([Jacod, Li, Mykland, Podolskij & Vetter 2009](#)).

The multivariate extension of realized variance, known as realized covariance, was first introduced to econometrics in [Andersen, Bollerslev, Diebold & Ebens \(2001\)](#) and the asymptotic

theory was first studied in [Barndorff-Nielsen & Shephard \(2004\)](#). In the absence of market microstructure noise, and when prices are synchronously observed, realized covariance estimates the quadratic covariation of prices. Estimators that are robust to microstructure noise and non-synchronous trading include multivariate realized kernels ([Barndorff-Nielsen, Hansen, Lunde & Shephard 2011](#)) and pre-averaging estimators ([Christensen, Kinnebrock & Podolskij 2010](#)) and realized QMLE ([Shephard & Xiu 2013](#)). Like their low-frequency counterparts, multivariate realized measures can be transformed to estimate other quantities, such as realized correlation or realized factor loadings (also known as realized beta). [Bollerslev & Zhang \(2003\)](#) employ realized factor loadings constructed in the Fama–French three-factor model to improve asset pricing predictions. [Barndorff-Nielsen & Shephard \(2004\)](#) derive the asymptotic distribution of realized betas. [Bandi & Russell \(2005\)](#) study the finite-sample properties of realized betas in the presence of market microstructure noise. [Andersen, Bollerslev, Diebold & Wu \(2006\)](#) find that realized betas are less persistent than realized variances and covariances and suggest modeling them as short-memory processes, and [Patton & Verardo \(2012\)](#) study the effect of earnings announcement on realized betas.

The recent financial crisis has motivated a growing literature on the measurement and management of systemic risk, often regarded as the risk measure conditional on the assumed market-wide scenario (for a survey we refer to [Bisias, Flood, Lo & Valavanis 2012](#)). Systemic risk measures can be designed rely exclusively on market-based data, which are widely available and are generally less stale than institution-reported measures. [Adrian & Brunnermeier \(2009\)](#) capture the particular institution’s time-varying marginal contribution to systemic risk as “CoVaR” defined as the Value-at-Risk (VaR) of the financial system conditional on a particular institution being under stress. [Acharya, Pedersen, Philippon & Richardson \(2012\)](#) propose a micro-founded model in which regulators levy a tax on each bank that optimizes its risk taking activity. The tax payment depends on each bank’s expected shortfall as well as its systemic expected shortfall which identifies its contribution to systemic risk. The systemic expected shortfall, defined as the expected amount a bank’s equity drops below its target level conditional on the systemic distress, increases with both leverage and marginal expected shortfall (MES). [Brownlees & Engle \(2012\)](#) implement an estimator of MES by using a TGARCH-cDCC model of time-varying covariances, and estimate the joint distribution of market and firm returns using a nonparametric approach.

In this paper, we introduce Factor HEAVY models, a new class of multivariate volatility models exploiting high frequency data and utilizing a factor approach to facilitate estimation in empirically relevant scenarios. . We exploit a factor decomposition of asset prices to build a model which is feasible in high-dimensions and estimable with an imbalanced panel. Our model resembles β -GARCH, which models the factor variance, conditional β and idiosyncratic variance

each with a GARCH-type evolution [Braun et al. \(1995a\)](#) model returns on individual assets as related through a common factor, so that $r_{i,t} = \beta_{i,t} r_{f,t} + \epsilon_{i,t}$, so that the covariance dynamics in the β -GARCH model are given by

$$\begin{aligned}\sigma_{f,t}^2 &= \theta_0 + \theta_1 r_{f,t-1}^2 + \theta_2 \sigma_{f,t-1}^2 \\ \beta_{i,t} &= \delta_{i,0} + \delta_{i,1} \frac{r_{f,t-1} r_{i,t-1}}{\sigma_{f,t-1}^2} + \delta_{i,2} \beta_{i,t-1} \\ \sigma_{i,t}^2 &= \alpha_{i,0} + \alpha_{i,1} (r_{i,t-1} - \beta_{i,t-1} r_{f,t-1})^2 + \alpha_{i,2} \sigma_{i,t-1}^2\end{aligned}\tag{1}$$

where $\sigma_{f,t}^2$ is the variance of the factor, or in a conditional CAPM, the market return. This specification uses natural proxies of the left-hand-side variables to act as shocks – squared returns of the factor for the factor variance, standardized cross-products for the β dynamics and squares of the innovation for the idiosyncratic variance.

We contribute to a growing literature that combines multivariate GARCH-type dynamics and realized measures, and explicitly model the closing return, not just the intra-daily return.¹ In multivariate HEAVY models ([Noureldin et al. 2012](#)), the conditional covariance is modeled as a smoothed function of recent lags of realized covariance matrix. [Maheu & Jin \(2013\)](#) model daily returns using a Wishart-autoregressive-like structure and propose a Bayesian estimation method. [Hansen, Lunde & Voev \(2014\)](#) propose the realized β GARCH model in which volatility and covolatility are separated modeled. In our models, daily returns on individual assets are driven by common factors with time-varying factor loadings.² The dynamics of factor volatility, conditional factor loadings and idiosyncratic volatility all follow the HEAVY structure where realized measures drive the dynamics of daily covariance ([Shephard & Sheppard 2010](#)). Factor HEAVY models have two advantages over multivariate HEAVY models. First, multivariate HEAVY models are directly parametrized on variances and covariance and so specify common dynamics for all second moments of asset returns. Second, multivariate HEAVY models suffer from the curse of dimensionality – not only in terms of the number of parameters in the model, but also in the dimension of the realized measure required to drive the dynamics. For example, when modeling the conditional covariance of 50 assets, a 50-dimensional realized covariance is required

¹Initial models that included realized measures focused primarily realized measure only modeling ([Halbleib & Voev \(2011\)](#), [Golosnoy, Gribisch & Liesenfeld \(2012\)](#)). While the persistence of realized measures is an interesting topic, it is not appropriate for most applications in portfolio allocation or risk-management.

²The time-variation in conditional β s has been debated in the literature for the last two decades. [Braun et al. \(1995b\)](#) use bivariate EGARCH models to find weak evidence of time-varying conditional β s. [Ferson & Harvey \(1993\)](#), [Bali & Engle \(2010\)](#), and [Hansen et al. \(2012\)](#) find significant time-series variation in the conditional β s. [Bali et al. \(2012\)](#) show the substantial time-varying conditional β s in the cross-section of daily stock returns.

each day. Our model only requires estimating low-dimensional realized measures irrespective of the number of assets in the model, and so can easily scale to empirically relevant dimensions. In the empirical analysis, we show that Factor HEAVY dominates other competing models in terms of in-sample performance. We also compare the out-of-sample ability of Factor HEAVY and the cDCC GARCH models (Aielli 2013) when forecasting volatility and covolatility, β s and MES. The results show that Factor HEAVY outperforms cDCC models, and that the gains are particularly substantial in short term forecasting. This superior performance at short horizons is particularly useful from a regulatory point-of-view since accurate and timely detection of changes in the covariance structure of returns is required when considering interventions.

The remainder of the paper is structured as follows. Section 2 introduces the Factor HEAVY models and discusses their properties. We initially focus the exposition on the 1-factor version of the model, although the extension to more factors is straight-forward. Section 3 discusses estimation and asymptotic properties. Section 4 describes the data used in the paper. Section 5 presents the result of empirical analysis including parameter estimation, in-sample performance, out-of-sample forecasting, an example of two-factor models and the robustness analysis. Section 6 concludes the paper.

2 Factor HEAVY Models

2.1 Notation and Model Setup

To facilitate the exposition of the model, the initial focus is on a 1-factor specification. The full K -factor specification is presented in section 2.7. Let $r_t = (r_{f,t}, r_{1,t}, r_{2,t}, \dots, r_{N,t})'$, denote a $N + 1$ by 1 vector of low-frequency, typically daily, returns. The first return is a pervasive factor and the remaining N are returns on individual assets assumed to be related through the factors.³ We denote the information set formed by the history of low-frequency returns with $\mathcal{F}_t^{LF}$, the natural filtration containing all past low-frequency returns. In the standard multivariate ARCH literature, returns are typically assumed to be conditionally normally distributed so that

$$r_t | \mathcal{F}_{t-1}^{LF} \sim N(0, \Omega_t).$$

Our interest is in modeling the conditional covariance of low-frequency returns using high-frequency derived realized measures, and so augment the information set with a realized measure that esti-

³While mixing intra-daily information and daily information is the natural application of the factor HEAVY models, this class of models is also useful for building models that mix daily (high frequency) and weekly or monthly (low frequency) data.

mates the quadratic covariation of the factor and individual assets, and denote the time t value of this $N+1$ by $N+1$ matrix-valued random variable RM_t . The realized measure could be a realized covariance or a more sophisticated noise-robust measure such as a realized kernel (Barndorff-Nielsen et al. 2011), pre-averaged realized variance (Christensen et al. 2010) or realized QMLE covariance (Shephard & Xiu 2013).⁴ We use $\mathcal{F}_t^{HF}$ to denote the filtration that contains both the realized measures and low-frequency returns, so that $\mathcal{F}_t^{LF} \subset \mathcal{F}_t^{HF}$.

We assume that returns, conditionally on the high-frequency information set, are normally distributed

$$r_t | \mathcal{F}_{t-1}^{HF} \sim N(0, \Sigma_t).$$

However, since the model now contains both high- and low-frequency data, it is necessary to specify a model for the process generating the realized measures. We assume that the realized measure is conditionally distributed as a Wishart with ν degrees of freedom, so that

$$RM_t | \mathcal{F}_{t-1}^{HF} \sim W_{N+1}(\nu, M_t/\nu)$$

where M_t is a positive definite matrix so that $RM_t = M_t^{1/2} \Xi_t (M_t^{1/2})'$ where $M_t^{1/2}$ is a matrix square-root of M_t and $\Xi_t | \mathcal{F}_{t-1}^{HF} \sim W_{N+1}(\nu, \frac{I_{N+1}}{\nu})$. We use a partitioning of the realized measures so that

$$RM_t = \begin{bmatrix} RM_{f,t} & RM_{f1,t} & \dots & RM_{fN,t} \\ RM_{1f,t} & RM_{11,t} & \dots & RM_{1N,t} \\ \vdots & \vdots & \ddots & \vdots \\ RM_{Nf,t} & RM_{N1,t} & \dots & RM_{NN,t} \end{bmatrix} \quad (2)$$

where $RM_{f,t}$ is a scalar measuring the quadratic variation of the factor, $RM_{ii,t}$ is a scalar realized measure estimating the quadratic variation of the individual asset and $RM_{if,t}$ is a scalar measuring the quadratic covariation between the asset and the factors.

We propose to model the conditional covariance of returns, Σ_t , as well as the conditional expectation of the realized measure, M_t , using a factor structure. The conditional covariance of

⁴While the choice of realized measure does not affect the description of the model, the steps required to estimate the model may differ depending on the estimator used. These issues are discussed in section 3.

returns and the conditional expectation of the realized measures can be written

$$\Sigma_t = \begin{bmatrix} \sigma_{f,t}^2 & \beta_{1,t}\sigma_{f,t}^2 & \cdots & \beta_{N,t}\sigma_{f,t}^2 \\ \beta_{1,t}\sigma_{f,t}^2 & \beta_{1,t}^2\sigma_{f,t}^2 + \sigma_{1,t}^2 & \cdots & \beta_{1,t}\beta_{N,t}\sigma_{f,t}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{N,t}\sigma_{f,t}^2 & \beta_{N,t}\beta_{1,t}\sigma_{f,t}^2 & \cdots & \beta_{N,t}^2\sigma_{f,t}^2 + \sigma_{N,t}^2 \end{bmatrix}, \quad (3)$$

and

$$M_t = \begin{bmatrix} \mu_{f,t} & \lambda_{1,t}\mu_{f,t} & \cdots & \lambda_{N,t}\mu_{f,t} \\ \lambda_{1,t}\mu_{f,t} & \lambda_{1,t}^2\mu_{f,t} + \mu_{1,t} & \cdots & \lambda_{1,t}\lambda_{N,t}\mu_{f,t} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{N,t}\mu_{f,t} & \lambda_{N,t}\lambda_{1,t}\mu_{f,t} & \cdots & \lambda_{N,t}^2\mu_{f,t} + \mu_{N,t} \end{bmatrix} \quad (4)$$

The model then requires specifications for the dynamics of the factor variance, $\sigma_{f,t}^2$ (and the corresponding value for the realized measure $\mu_{f,t}$), the factor loadings $\beta_{i,t}$ ($\lambda_{i,t}$) and the idiosyncratic variances $\sigma_{i,t}^2$ ($\mu_{i,t}$). Throughout the paper, we use the notation $E_{t-1}(\cdot) = E(\cdot \mid \mathcal{F}_{t-1}^{HF})$ used to denote the expectation conditional on the high frequency information set. The Wishart distribution is in the linear exponential family and is the natural candidate for positive definite matrix-valued shocks. Moreover, the transformations required when building factor-based models are known in closed form and lead to simple expressions for conditional expectations of the quantities of interest.

Proposition 1. *Under the assumption $RM_t \mid \mathcal{F}_{t-1}^{HF} \sim W_{N+1}(\nu, \frac{M_t}{\nu})$, the conditional expectation of realized β , $E_{t-1}[R\beta_{i,t}] = \lambda_{i,t}$ and that of realized idiosyncratic variance $E_{t-1}[RV_{i,t}] = \frac{\nu-1}{\nu}\mu_{i,t}$.*

2.2 Factor Variance Dynamics

The factor dynamics are assumed to follow a HEAVY structure. Suppose there is a single factor, in which case the factor returns are assumed to be conditionally normal with mean zero,

$$r_{f,t} = \sigma_{f,t}\xi_{f,t}$$

where the conditional variance of the factor return is measurable with respect to $\mathcal{F}_{t-1}^{HF}$ and $E_{t-1}[\xi_{f,t}^2] = 1$. The conditional variance of the factor is driven by the realized measure corresponding to the

quadratic variation of the factor, and so

$$\sigma_{f,t}^2 = \theta_0 + \theta_1 RM_{f,t-1} + \theta_2 \sigma_{f,t-1}^2. \quad (5)$$

This is not a complete model and only allows for 1-step forecasts, and so we complete the model by specifying a model for the realized measure. The dynamics of the conditional mean of the realized measure follow a similar process to the factor conditional variance,

$$\mu_{f,t} = \theta_0^M + \theta_1^M RM_{f,t-1} + \theta_2^M \mu_{f,t-1}. \quad (6)$$

These two equations correspond to the HEAVY model of (Shephard & Sheppard 2010).

2.3 Factor Loading Dynamics

Returns on individual assets are related through exposure to a common factor through a time varying loading. The factor loading of asset i is

$$\beta_{i,t} = \frac{\sigma_{if,t}}{\sigma_{f,t}^2} = \frac{\text{Cov}[r_{f,t}, r_{i,t}]}{\text{V}[r_{f,t}]}.$$

Factor loadings are also driven by realized measures, where the natural analogue of the β , commonly referred to as the realized β is used,

$$R\beta_{i,t} = \frac{RM_{if,t}}{RM_{f,t}}.$$

Without microstructure noise, Barndorff-Nielsen & Shephard (2004) show that when the realized measure is realized covariance, then the realized β is a consistent estimator of the ratio between integrated equity covariance with the factor and integrated variance of the factor return. The dynamic factor loadings evolve as

$$\beta_{i,t} = \delta_{i,0} + \delta_{i,1} R\beta_{i,t-1} + \delta_{i,2} \beta_{i,t-1} \quad (7)$$

and we complete the model by specifying dynamics for the realized β . Let $\lambda_{i,t} = E_{t-1}[R\beta_{i,t}]$, then

$$\lambda_{i,t} = \delta_{i,0}^M + \delta_{i,1}^M R\beta_{i,t-1} + \delta_{i,2}^M \lambda_{i,t-1}. \quad (8)$$

There is no observation equation for the realized β and its measurement occurs jointly with the idiosyncratic variances.

2.4 Idiosyncratic Dynamics

The final component of the model is the idiosyncratic variance, defined as $E_{t-1} \left[(r_{i,t} - \beta_{i,t} r_{f,t})^2 \right]$. Since the factor and the individual asset return are assumed to be conditionally normally, we can define the idiosyncratic shock as

$$\epsilon_{i,t} = r_{i,t} - \beta_{i,t} r_{f,t}$$

where $\epsilon_{i,t} = \sigma_{i,t} \xi_{i,t}$ and $\sigma_{i,t}^2$ is the conditional variance of $r_{i,t}$ given the factor return. By construction, the N by 1 vector $\epsilon_t = (\epsilon_{1,t}, \dots, \epsilon_{N,t})'$ is contemporaneously uncorrelated, and we further assume that there are no volatility spill-overs between assets. We use the realized idiosyncratic variance, defined

$$RIV_{i,t} = RM_{ii,t} - (R\beta_{i,t})^2 RM_{f,t},$$

to drive the dynamics of the idiosyncratic variance. We assume a univariate HEAVY-like structure for the idiosyncratic variance, so that

$$\sigma_{i,t}^2 = \alpha_{i,0} + \alpha_{i,1} RIV_{i,t-1} + \alpha_{i,2} \sigma_{i,t-1}^2. \quad (9)$$

The model is completed by specifying the evolution of the realized idiosyncratic volatility as

$$\mu_{i,t} = \alpha_{i,0}^M + \alpha_{i,1}^M RIV_{i,t-1} + \alpha_{i,2}^M \mu_{i,t-1}. \quad (10)$$

2.5 Stationarity

The model specification contains two distinct components. The first component, which determines the dynamics of the conditional covariance of the low-frequency data, is collectively referred to as the HEAVY-P equations,

$$\begin{aligned} \sigma_{f,t}^2 &= \theta_0 + \theta_1 RM_{f,t-1} + \theta_2 \sigma_{f,t-1}^2 \\ \beta_{i,t} &= \delta_{i,0} + \delta_{i,1} R\beta_{i,t-1} + \delta_{i,2} \beta_{i,t-1} \\ \sigma_{i,t}^2 &= \alpha_{i,0} + \alpha_{i,1} RIV_{i,t} + \alpha_{i,2} \sigma_{i,t-1}^2. \end{aligned} \quad (11)$$

where the final two equations are repeated N times. We refer to the second component, which governs the dynamics of the realized measure, as the HEAVY-M equations,

$$\begin{aligned}
\mu_{f,t} &= \theta_0^M + \theta_1^M RM_{f,t-1} + \theta_2^M \mu_{f,t-1} \\
\lambda_{i,t} &= \delta_{i,0}^M + \delta_{i,1}^M R\beta_{i,t-1} + \delta_{i,2}^M \lambda_{i,t-1} \\
\mu_{i,t} &= \alpha_{i,0}^M + \alpha_{i,1}^M RIV_{i,t-1} + \alpha_{i,2}^M \mu_{i,t-1}
\end{aligned} \tag{12}$$

where the final two equations are also repeated for each individual asset. This structure leads to a relatively simple condition to determine whether the model is covariance stationary since both the HEAVY-P and the HEAVY-M components are driven exclusively by the realized measures. In particular, the coefficients on the realized measures in the HEAVY-P play no role in the stationarity since these simply act as sensitivities and there is no feed-back from the low-frequency observations into the realized measures.

From the properties of the Wishart distribution, all of the elements can be decomposed into their conditional expectation and a white noise error, so that $RM_{f,t} = \mu_{f,t} + \tilde{\eta}_{f,t}$, $R\beta_{i,t} = \lambda_{i,t} + \tilde{\tau}_{i,t}$, and $RIV_{i,t} = \frac{\nu-1}{\nu}\mu_{i,t} + \tilde{l}_{i,t}$. Letting $\nu_t = (0, 0, 0, \tilde{\eta}_{f,t}, \tilde{\tau}_{i,t}, \tilde{l}_{i,t})'$, a mean-zero white-noise vector, the standard Factor HEAVY have a VARMA(1,1) representation. Specifically,

$$\tilde{a}_t = c + B\tilde{a}_{t-1} + \nu_t - D\nu_{t-1} \tag{13}$$

where

$$\tilde{a}_t = \begin{bmatrix} \sigma_{f,t}^2 \\ \beta_{i,t} \\ \sigma_{i,t}^2 \\ RM_{f,t} \\ R\beta_{i,t} \\ RIV_{i,t} \end{bmatrix}, c = \begin{bmatrix} \theta_0 \\ \delta_{i,0} \\ \alpha_{i,0} \\ \theta_0^M \\ \delta_{i,0}^M \\ \frac{\nu-1}{\nu}\alpha_{i,0}^M \end{bmatrix}, D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \theta_2^M & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_{i,2}^M & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha_{i,2}^M \end{bmatrix}$$

and

$$B = \begin{bmatrix} \theta_2 & 0 & 0 & \theta_1 & 0 & 0 \\ 0 & \delta_{i,2} & 0 & 0 & \delta_1 & 0 \\ 0 & 0 & \alpha_{i,2} & 0 & 0 & \alpha_1 \\ 0 & 0 & 0 & \theta_1^M + \theta_2^M & 0 & 0 \\ 0 & 0 & 0 & 0 & \delta_{i,1}^M + \delta_{i,2}^M & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{\nu-1}{\nu}\alpha_{i,1}^M + \alpha_{i,2}^M \end{bmatrix}.$$

The covariance stationarity for VARMA(1,1) requires that the eigenvalues of B are all less than one in modulus, which we formalize in the following proposition.

Proposition 2. *The standard Factor HEAVY with a single factor is covariance stationary if*

$$0 \leq \theta_2 < 1, 0 \leq \delta_{i,2} < 1, 0 \leq \alpha_{i,2} < 1, 0 \leq \theta_1^M + \theta_2^M < 1, 0 \leq \delta_{i,1}^M + \delta_{i,2}^M < 1, \\ 0 \leq \frac{\nu-1}{\nu} \alpha_{i,1}^M + \alpha_{i,2}^M < 1.$$

2.6 Forecasting

We focus on forecasting conditional factor variances, β s and idiosyncratic variances. One-step forecasts are directly given in the HEAVY-P equations. To compute multi-step forecasts, we utilize the recursive structure of the HEAVY-M equations.

Proposition 3. *The s -step forecasts in the Factor HEAVY model is given by*

$$\sigma_{f,t+s|t}^2 = E_t(\sigma_{f,t+s}^2) = \theta_2^{s-1} \sigma_{f,t+1}^2 + \theta_0 \frac{1 - \theta_2^{s-1}}{1 - \theta_2} \quad (14)$$

$$+ \theta_1 \sum_{i=1}^{s-1} \theta_2^{i-1} (\theta_0^M \frac{1 - (\theta_1^M + \theta_2^M)^{s-1-i}}{1 - (\theta_1^M + \theta_2^M)} + (\theta_1^M + \theta_2^M)^{s-1-i} \mu_{f,t+1})$$

$$\beta_{i,t+s|t} = E_t(\beta_{i,t+s}) = \delta_{i,2}^{s-1} \beta_{i,t+1} + \delta_{i,0} \frac{1 - \delta_{i,2}^{s-1}}{1 - \delta_{i,2}} \\ + \delta_{i,1} \sum_{i=1}^{s-1} \delta_{i,2}^{i-1} (\delta_{i,0}^M \frac{1 - (\delta_{i,1}^M + \delta_{i,2}^M)^{s-1-i}}{1 - (\delta_{i,1}^M + \delta_{i,2}^M)} + (\delta_{i,1}^M + \delta_{i,2}^M)^{s-1-i} \lambda_{i,t+1})$$

$$\sigma_{i,t+s|t}^2 = E_t(\sigma_{i,t+s}^2) = \alpha_{i,2}^{s-1} \sigma_{i,t+1}^2 + \alpha_{i,0} \frac{1 - \alpha_{i,2}^{s-1}}{1 - \alpha_{i,2}} \\ + \alpha_{i,1} \sum_{i=1}^{s-1} \alpha_{i,2}^{i-1} (\frac{\nu-1}{\nu} \alpha_{i,0}^M \frac{1 - (\frac{\nu-1}{\nu} \alpha_{i,1}^M + \alpha_{i,2}^M)^{s-1-i}}{1 - (\frac{\nu-1}{\nu} \alpha_{i,1}^M + \alpha_{i,2}^M)} + (\frac{\nu-1}{\nu} \alpha_{i,1}^M + \alpha_{i,2}^M)^{s-1-i} \mu_{i,t+1})$$

The proof is presented in the appendix. While these are the forecasts for the components of the conditional covariance, they do not lead to direct forecasts of the conditional covariance. An unbiased estimate could be computed using simulation, although we find that this error is sufficiently small so that simulation is not warranted.

2.7 Multiple Factors

The 1-factor model can be directly extended to include multiple factors. In the multi-factor HEAVY model, returns on individual assets are determined by K factors and an idiosyncratic shock,

$$r_{i,t} = \sum_{k=1}^K \beta_{i,k,t} r_{fk,t} + \epsilon_{i,t} = \sum_{k=1}^K \beta_{i,k,t} \sigma_{fk,t} e_{fk,t} + \sigma_{i,t} e_{i,t}$$

in which $r_{fk,t}$ and $\sigma_{fk,t}$ denote the daily return on the k^{th} factor and its volatility respectively; $\beta_{i,k,t}$ represents the conditional loading of $r_{i,t}$ with respect to $r_{fk,t}$; $\{e_{fk,t}\}$ is an i.i.d. innovation sequence with zero mean and unit variance. $\epsilon_{i,t}$ is uncorrelated with each factor return given $\mathcal{F}_{t-1}^{\text{HF}}$.

We now assume that the return vector r_t contains K factors and N individual assets where the K factors are in the first K positions, and the realized measure is $K + N$ by $K + N$ where the upper K by K block measures the quadratic covariation of the factors. We use block-partitions of the conditional covariance,

$$\Sigma_t = \begin{bmatrix} \Sigma_{f,t} & \Sigma_{f1,t} & \dots & \Sigma_{fN,t} \\ \Sigma_{1f,t} & \Sigma_{11,t} & \dots & \Sigma_{1N,t} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{Nf,t} & \Sigma_{N1,t} & \dots & \Sigma_{NN,t} \end{bmatrix},$$

and the realized measure,

$$RM_t = \begin{bmatrix} RM_{f,t} & RM_{f1,t} & \dots & RM_{fN,t} \\ RM_{1f,t} & RM_{11,t} & \dots & RM_{1N,t} \\ \vdots & \vdots & \ddots & \vdots \\ RM_{Nf,t} & RM_{N1,t} & \dots & RM_{NN,t} \end{bmatrix}$$

where the upper left block is K by K , the if blocks are 1 by K and the ii blocks are scalars. We allow for a general dependence structure of factors, and so the conditional loadings for asset i are

$$\beta_{i,t} = \Sigma_{f,t}^{-1} \Sigma_{f1,t}.$$

Similarly, define the multi-factor realized β s for asset i as

$$R\beta_{i,t} = RM_{f,t}^{-1} RM_{fi,t}$$

where $RM_{f,t}$ is K by K and the $RM_{fi,t}$ is K by 1, and the realized idiosyncratic variance as

$$RIV_{i,t} = RM_{ii,t} - R\beta'_{i,t}RM_{f,t}R\beta_{i,t}.$$

In the K -factor HEAVY model, the multiple factors follow a multivariate HEAVY model (Noureldin et al. 2012), while each of the factor loadings and idiosyncratic volatility follow the same dynamics as in the 1-factor model. The HEAVY-P equations are then

$$\begin{aligned}\Sigma_{f,t} &= CC' + ARM_{f,t-1}A' + B\Sigma_{f,t-1}B' \\ \beta_{i,t} &= \delta_{i,0} + \delta_{i,1}R\beta_{i,t-1} + \delta_{i,2}\beta_{i,t-1} \\ \sigma_{i,t}^2 &= \alpha_{i,0} + \alpha_{i,1}RIV_{i,t-1} + \alpha_{i,2}\sigma_{i,t-1}^2\end{aligned}\tag{15}$$

where $\beta_{i,t} = (\beta_{i,1,t}, \dots, \beta_{i,K,t})$, $\delta_{i,0}$ is a K by 1 vector, $\delta_{i,1}$ and $\delta_{i,2}$ are diagonal matrices and A, B and C are parameter matrices which satisfy the assumptions given in Noureldin et al. (2012). The HEAVY-M equations are similarly modified, and are

$$\begin{aligned}M_{f,t} &= C^M (C^M)' + A^M RM_{f,t-1} (A^M)' + B^M M_{f,t-1} (B^M)' \\ \lambda_{i,t} &= \delta_{i,0}^M + \delta_{i,1,k}^M R\beta_{i,t-1} + \delta_{i,2}^M \lambda_{i,t-1} \\ \mu_{i,t} &= \alpha_{i,0}^M + \alpha_{i,1}^M RIV_{i,t-1} + \alpha_{i,2}^M \mu_{i,t-1}\end{aligned}\tag{16}$$

where $M_{f,t} = E_{t-1}(RM_{f,t})$. As before, the final two equations in (15) and (16) are repeated for each asset. Equations (15) and (16) constitute the K -factor HEAVY model. The condition for covariance stationarity of the full model is similar to that in the 1-factor model, and so in the interest of brevity, we omit a detailed discussion.

3 Estimation and Inference

3.1 Estimation

This discussion focuses on the 1-factor model, before turning to the K -factor models. The conditional likelihood of the returns and the realized measures is the natural method to estimate the parameters. The parameters in the HEAVY-P equations are estimated by conditional maximum

likelihood by maximizing

$$\operatorname{argmax}_{\psi} L = \sum_{t=1}^T l_t(\psi; r_t) \quad (17)$$

where $\psi = (\theta', \phi_1', \dots, \phi_N')'$, $\theta = (\theta_0, \theta_1, \theta_2)'$, $\phi_i = (\delta_0, \delta_1, \delta_2, \alpha_0, \alpha_1, \alpha_2)'$,

$$l_t(\psi; r_t) = -\frac{1}{2} (\ln |\Sigma_t| + r_t' \Sigma_t^{-1} r_t) + c \quad (18)$$

and c is a term which does not depend on the model parameters.⁵ The model structure can be directly exploited to simplify estimation by expressing the log-likelihood in two components – one which measures the likelihood of the common factor, and one which measures the likelihood of the idiosyncratic errors, $r_{i,t} - \beta_{i,t} r_{f,t}$.

Proposition 4. *The joint log-likelihood of the daily returns can be equivalently expressed*

$$l_t = \underbrace{-\frac{1}{2} \left(\ln(\sigma_{f,t}^2) + \frac{r_{f,t}^2}{\sigma_{f,t}^2} \right)}_{l_{f,t}: \text{factor}} + \sum_{i=1}^N \underbrace{-\frac{1}{2} \left(\ln(\sigma_{i,t}^2) + \frac{(r_{i,t} - \beta_{i,t} r_{f,t})^2}{\sigma_{i,t}^2} \right)}_{l_{i,t}: \text{idiosyncratic } i} + c \quad (19)$$

This decomposition leads to a natural 2-step estimator, where the parameters of the $l_{f,t}$ are first maximized, and then the parameters governing the conditional factor loadings and idiosyncratic volatilities are estimated.⁶

The parameters of the HEAVY-M equations are estimated by maximizing the standardized Wishart log-likelihood,

$$\operatorname{argmax}_{\psi^M} L^M = \sum_{t=1}^T l_t^M(\psi^M; RM_t)$$

where $\psi^M = [(\theta^M)', (\phi_1^M)', \dots, (\phi_N^M)']'$, $\theta^M = (\theta_0^M, \theta_1^M, \theta_2^M)'$, $\phi_i^M = (\delta_0^M, \delta_1^M, \delta_2^M, \alpha_0^M, \alpha_1^M, \alpha_2^M)'$,

$$l_t^M = -\frac{\nu}{2} (\ln |M_t| + \operatorname{tr}(M_t^{-1} (RM_t))) + c_\nu^M \quad (20)$$

and c_ν^M is a constant conditional on the shape parameter of the standardized Wishart, ν . Our interest is in the parameters of the dynamics, and so we do not estimate this parameter. Using

⁵The parameter dynamics are all recursive and so depend on values for observations 0. We use a backward exponentially weighted moving average based on the first $\tau = \lceil T^{0.25} \rceil$ observations of the realized measure to initialize the process so that $M_0 = \sum_{t=1}^{\tau} w_t M_t$ where $w_t = .06(.94)^{t-1} / \sum_{j=1}^{\tau} .06(.94)^{j-1}$.

⁶The second step of the estimation process involves N optimizations, although we refer to this as a single step since the ordering of the assets does not matter.

the structure of M_t , this log-likelihood can be similarly decomposed into two components,

$$l_t^M = \underbrace{-\frac{\nu}{2} \left(\ln(\mu_{f,t}) + \frac{RM_{f,t}}{\mu_{f,t}} \right)}_{l_{f,t}^M: \text{factor}} + \sum_{i=1}^N \underbrace{-\frac{\nu}{2} \left(\ln(\mu_{i,t}) + \frac{(\lambda_{i,t}^2 RM_{f,t} - 2\lambda_{i,t} RM_{fi,t} + RM_{ii,t})}{\mu_{i,t}} \right)}_{l_{i,t}^M: \text{idiosyncratic } i} + c_\nu^M. \quad (21)$$

This decomposition also leads to a natural 2-step estimation strategy, where the parameters governing the factor dynamics are first estimated and then the parameters of the idiosyncratic volatility are estimated. This decomposition comes directly from the model and does not require correct specification or other restrictions on the realized measure. This is particularly useful since estimation of the model parameters only requires storing the realized measure for the factor, the assets and between the factor and assets, and not the complete $N + 1$ by $N + 1$ realized measure. The likelihood structure is particularly useful when using noise-robust realized measures which are known to suffer from data attrition due to refresh-time sampling when the number of assets, N , is large. An alternative to using a single realized measure for all assets is to construct N estimators including the factor and each asset. This would allow the maximum amount of data for a pair to be used without loss. We discuss additional issues in using pairwise noise-robust estimators in section 4.1.1.

The joint likelihoods of the K -factor model can be similarly decomposed, so that

$$l_t = \underbrace{-\frac{1}{2} \left(\ln |\Sigma_{f,t}| + r'_{f,t} \Sigma_{f,t}^{-1} r_{f,t} \right)}_{l_{f,t}: \text{factor}} + \sum_{i=1}^N \underbrace{-\frac{1}{2} \left(\ln(\sigma_{i,t}^2) + \frac{\left(r_{i,t} - \sum_{k=1}^K \beta_{i,k,t} r_{fk,t} \right)^2}{\sigma_{i,t}^2} \right)}_{l_{i,t}: \text{idiosyncratic } i} + c$$

and

$$l_t^M = \underbrace{-\frac{\nu}{2} \left(\ln |M_{f,t}| - \text{tr} \left(M_{f,t}^{-1} (RM_{f,t}) \right) \right)}_{l_{f,t}^M: \text{factor}} + \sum_{i=1}^N \underbrace{-\frac{\nu}{2} \left(\ln(\mu_{i,t}) - \frac{\lambda'_{i,t} RM_{f,t} \lambda_{i,t} - 2\lambda'_{i,t} RM_{fi,t} + RM_{ii,t}}{\mu_{i,t}} \right)}_{l_{i,t}^M: \text{idiosyncratic } i} + c_\nu^M$$

where $\lambda_{i,t} = (\lambda_{i,1,t}, \dots, \lambda_{i,K,t})'$. The factor-HEAVY structure preserves the variation-free nature of the HEAVY framework in the sense that the lagged terms from one set of equations do not appear

in the other.

3.2 Quasi-likelihood Based Asymptotic Inference

Similar to the HEAVY models, the parameter estimators in factor HEAVY equations have the common asymptotic properties of quasi-maximum likelihood estimators. Since the parameters have no link between factor and equity equations, and between HEAVY-P and HEAVY-M equations, we can consider their asymptotic properties separately. Here we focus on the 1-factor HEAVY model for simplicity. The score equations for the corresponding likelihood, evaluated at the estimated parameters are

$$\begin{aligned}\sum_{t=1}^T S_{f,t}(\hat{\theta}) &= 0 & \sum_{t=1}^T S_{f,t}^M(\widehat{\theta^M}) &= 0 \\ \sum_{t=1}^T S_{i,t}(\hat{\phi}_i) &= 0 & \sum_{t=1}^T S_{i,t}^M(\widehat{\phi_i^M}) &= 0\end{aligned}$$

where $S_{f,t}(\theta) = \partial l_{f,t}/\partial \theta$, $S_{i,t}(\phi_i) = \partial l_{i,t}/\partial \phi_i$, $S_{f,t}^M(\theta^M) = \partial l_{f,t}^M/\partial(\theta^M)$, $S_{i,t}^M(\phi_i^M) = \partial l_{i,t}^M/\partial(\phi_i^M)$. Let θ_o , $\phi_{i,o}$, θ_o^M and $\phi_{i,o}^M$ indicate the true parameter values. The scores evaluated at these values are martingale difference sequences with respect to $\mathcal{F}_{t-1}^{\text{HF}}$. Under certain regularity conditions (see [Bollerslev & Wooldridge \(1992\)](#), [Newey & McFadden \(1994\)](#), [White \(1996\)](#), *inter alia.*), we have

$$\begin{aligned}\sqrt{T}(\hat{\theta} - \theta_o) &\rightarrow N(0, \mathcal{I}_\theta^{-1} \mathcal{J}_\theta \mathcal{I}_\theta^{-1}) & \sqrt{T}(\widehat{\theta^M} - \theta_o^M) &\rightarrow N(0, (\mathcal{I}_{\theta^M}^M)^{-1} \mathcal{J}_{\theta^M}^M (\mathcal{I}_{\theta^M}^M)^{-1}) \\ \sqrt{T}(\hat{\phi}_i - \phi_{i,o}) &\rightarrow N(0, \mathcal{I}_{\phi_i}^{-1} \mathcal{J}_{\phi_i} \mathcal{I}_{\phi_i}^{-1}) & \sqrt{T}(\widehat{\phi_i^M} - \phi_{i,o}^M) &\rightarrow N(0, (\mathcal{I}_{\phi_i^M}^M)^{-1} \mathcal{J}_{\phi_i^M}^M (\mathcal{I}_{\phi_i^M}^M)^{-1})\end{aligned}$$

where

$$\begin{aligned}\mathcal{I}_\theta &= -\frac{1}{T} \sum_{t=1}^T \mathbb{E} \left(\frac{\partial S_{f,t}(\theta)}{\partial \theta'} \bigg|_{\theta=\theta_o} \right), \quad \mathcal{J}_\theta = \text{avar} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T S_{f,t}(\theta_o) \right) \\ \mathcal{I}_{\phi_i} &= -\frac{1}{T} \sum_{t=1}^T \mathbb{E} \left(\frac{\partial S_{i,t}(\phi_i)}{\partial \phi_i'} \bigg|_{\phi_i=\phi_{i,o}} \right), \quad \mathcal{J}_{\phi_i} = \text{avar} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T S_{i,t}(\phi_{i,o}) \right) \\ \mathcal{I}_{\theta^M}^M &= -\frac{1}{T} \sum_{t=1}^T \mathbb{E} \left(\frac{\partial S_{f,t}^M(\theta^M)}{\partial (\theta^M)'} \bigg|_{\theta^M=\theta_o^M} \right), \quad \mathcal{J}_{\theta^M}^M = \text{avar} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T S_{f,t}^M(\theta_o^M) \right)\end{aligned}$$

$$\mathcal{I}_{\phi_i^M}^M = -\frac{1}{T} \sum_{t=1}^T \mathbb{E} \left(\frac{\partial S_{i,t}^M(\phi_i^M)}{\partial (\phi_i^M)'} \bigg|_{\phi_i^M = \phi_{i,o}^M} \right), \quad \mathcal{J}_{\phi_i^M}^M = \text{avar} \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T S_{i,t}^M(\phi_{i,o}^M) \right)$$

We have omitted calculation of the cross-terms in the variance covariance between components of the model. While these are generally not of interest, calculation of these is straight-forward. Inference in the factor HEAVY model is simplified when compared to typical multivariate volatility models since the estimating equations have a natural block structure due to the factor structure.

4 Empirical Analysis and Model Evaluation

4.1 Data and Descriptive Statistics

We apply the model to a sample containing 40 large U.S. financial firms from July 1, 2000 to June 30, 2010 so that a maximum of 2,511 daily observations are available. The panel is slightly heterogeneous since some of the included firms stop trading during the financial crisis. High frequency quote and trade data are extracted from the Trade and Quote (TAQ) database, and daily price data are extracted from Center for Research in Security Prices (CRSP) stock database. High-frequency data typically contains mis-recordings and other erroneous data, and so all price data was cleaned using a slightly modified set of rules to those used in [Barndorff-Nielsen et al. \(2009\)](#).⁷ Our initial focus is on a 1-factor model similar to a conditional CAPM ([Jagannathan & Wang 1996](#), [Lewellen & Nagel 2006](#)). We chose the S&P 500 as the market proxy, and use the SPDR S&P 500 (SPY), a highly liquid ETF that tracks the S&P 500 index.

In most of the application we use realized covariance as the realized measure. We sample all prices using last-price interpolation and use a sparse sampling scheme. Our preferred estimator is based on 10-minute sampling with subsampling every minute. Suppose $p_{j,t}$ is the j^{th} log price vector on day t containing the log prices of the factor and the 40 firms. Then the sub-sampled realized covariance is defined

$$RC_t^{SS} = \frac{m-s}{s(m-s)} \sum_{i=1}^{m-s} \bar{r}_{i,t} \bar{r}_{i,t}'$$

where $\bar{r}_{i,t} = \sum_{j=i}^s r_{i+j,t}$, s is the length of the block (e.g. 10), $r_{j,t} = p_{j,t} - p_{j-1,t}$ are the high-frequency returns, m is the number of price samples (e.g. 390 when using 1-minute returns in U.S. equity data). We consider alternative specifications where we vary $s \in \{5, 10, 15, 30\}$, as

⁷The sole modification was to use the market each day that had the highest trading volume rather than always selecting NYSE.

well as a realized kernel where we use the non-flat Parzen kernel and the bandwidth selection procedure outlined in [Barndorff-Nielsen et al. \(2011\)](#).

Figure 1 shows annualized realized volatility of the factor, average annualized realized volatility of all firm equity returns, average realized correlations between SPY and equity returns, and average realized β s. It is obvious that from summer 2007, the beginning of the financial crisis, daily returns become more volatile than before the crisis. The returns on the factor are less volatile than those on equity returns. During the crisis, the average realized correlations increased relative to their pre-crisis values, although the changes in dependence are far more striking in terms of the average realized β s.

One of the primary difficulties encountered when incorporating realized measures into a model of the conditional covariance of low-frequency data is the missing overnight data. In models of the conditional variance, it is common to assume that the full-day variance can be simply scaled from the within trading-day variance. When studying multivariate quantities, there are more possibilities, and in particular the covariance and the variances may not scale with a common factor. We examine the choice of modeling space in Figure 2, where we compare the intra-daily and overnight β s and as well as the intra-daily and overnight correlations between the factor and equity returns. We computed these using only the open-to-close and close-to-open returns (not the other intra-daily data). The result shows β behaves very differently from correlation – most β s lie near the 45-degree line, indicating that β s are very stable throughout the entire day. The relationship between intra-daily and overnight correlations appears to be more complicated, and many of the firms have stronger correlations with the factor during the market opening hours than during the overnight period. This is consistent with systematic news arriving during market hours and idiosyncratic news arriving after the market closes. This difference indicates that models incorporating high-frequency data that parameterize the conditional β are better suited than models which use other transformations of the conditional covariance.

4.1.1 Data Transformations

We are primarily interested in the covariance of returns over the entire day, including both the period where markets are active and the overnight return. Realized measures can only be computed for the period where the market is actively traded, and so we use a transformation to ensure that the average value of the realized measure is the same as the average value of the outer product of returns. The modified realized measure is constructed as

$$\widetilde{RM}_t = \hat{\Lambda} RM_t \hat{\Lambda}'$$

Figure 1: Top Left: Annualized realized volatility of SPY, the square root of the mean of 252 times 1-minute realized variance; Top Right: average annualized realized volatility of all firm equity returns; Bottom Left: average realized correlations between SPY and firm equity returns; Bottom Right: average realized β s.

Figure 2: The left panel shows intra-daily and overnight market β s for the 40 firms in the application. The right panel shows the intra-daily and overnight correlations between the firm returns and the market for the same firms.

where $\hat{\Lambda} = \bar{\Sigma}^{1/2} \bar{M}^{-1/2}$, $\bar{\Sigma} = T^{-1} \sum_{t=1}^T r_t r_t'$ and $\bar{M} = T^{-1} \sum_{t=1}^T RM_t$. This transformation has some limitations, and in particular the matrix square-root of $\bar{\Sigma}$ may not be precisely estimated. Since our preferred estimator only uses the realized measure of the factor(s) and the individual assets – but not the realized measure between the individual assets – we can define a similar transformation using the $K + 1$ by $K + 1$ realized measures required by the estimator,

$$\widetilde{RM}_{i,t} = \hat{\Lambda}_i RM_{i,t} \hat{\Lambda}_i' \quad (22)$$

where $\hat{\Lambda}$ is estimated using only the K factors and asset i , and

$$RM_{i,t} = \begin{bmatrix} RM_{f,t} & RM_{fi,t} \\ RM_{if,t} & RM_{ii,t} \end{bmatrix}.$$

Finally, a particular difficulty arises when using noise-robust estimators constructed using only the K factors and the individual assets. This occurs since these estimators typically use some form of refresh time sampling and so will not produce identical estimates of the quadratic covariation of the factors. We propose to first transform these so that they have the same estimate of the quadratic covariation of the factors, and then these modified realized measures can be standardized using eq. (22). Begin by partitioning the individual realized measure

$$RM_{i,t} = \begin{bmatrix} RM_{f,t}^i & RM_{fi,t} \\ RM_{if,t} & RM_{ii,t} \end{bmatrix}.$$

where the notation $RM_{f,t}^i$ is used to indicate that this estimate of the quadratic covariation of the factors is specific to the noise-robust measure that includes asset i . Our solution is to use a common estimator of the quadratic covariation of the factors, and then to enforce that this common estimator is the same in all component realized measures. A natural estimator would be the noise robust estimator (e.g. realized kernel) applied only to the factors on day t . Denote that estimator $RM_{f,t}^C$ where the C denotes common. The modified estimators of quadratic covariation are then

$$\widehat{RM}_{i,t} = \Pi_{i,t} RM_{i,t} \Pi'_{i,t}$$

where

$$\Pi_{i,t} = \begin{bmatrix} \left(RM_{f,t}^C\right)^{1/2} \left(RM_{f,t}^i\right)^{-1/2} & 0 \\ 0 & 1 \end{bmatrix}.$$

When there is a single factor, the transformation matrix takes a simple form where

$$\begin{aligned} \widehat{RM}_{i,t} &= \begin{bmatrix} \sqrt{\frac{RM_{f,t}^C}{RM_{f,t}^i}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} RM_{f,t}^i & RM_{fi,t} \\ RM_{if,t} & RM_{ii,t} \end{bmatrix} \begin{bmatrix} \sqrt{\frac{RM_{f,t}^C}{RM_{f,t}^i}} & 0 \\ 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} RM_{f,t}^C & RM_{fi,t} \sqrt{\frac{RM_{f,t}^C}{RM_{f,t}^i}} \\ RM_{if,t} \sqrt{\frac{RM_{f,t}^C}{RM_{f,t}^i}} & RM_{ii,t} \end{bmatrix} \end{aligned}$$

Note that if a single realized measure is used across all assets, such as a single multivariate realized kernel (or realized covariance) to estimate the integrated covariance, then $\left(RM_{f,t}^C\right)^{1/2} \left(RM_{f,t}^i\right)^{-1/2} = I_K$ and the estimator is unmodified.

4.2 In-Sample Performance

The Factor HEAVY model was estimated using the complete sample available for each firm. Table 1 shows the parameter estimates for the factor and a representative firm, Capital One (COF). The estimates for the market volatility (θ and θ^M) are similar to those in Shephard & Sheppard (2010), and the model indicates that the factor volatility is highly persistent since $\theta_1^M + \theta_2^M \approx 1$, a well documented feature in the ARCH-literature. However, the responsiveness to volatility news, as measured by θ_1 , is much higher than is commonly found in low-frequency models. Conditional β s are also highly persistent, and somewhat less responsive to news. Idiosyncratic volatility is also extremely persistent, and is highly sensitive to recent idiosyncratic volatility news. All esti-

mates in the table are statistically significant, as are virtually all parameters in the model⁸ The parameters of the HEAVY-M equations indicate substantial persistence of all three series.

	Factor		β		Idiosyncratic	
HEAVY-P	θ_1	θ_2	$\delta_{i,1}$	$\delta_{i,2}$	$\alpha_{i,1}$	$\alpha_{i,2}$
	0.26 (0.05)	0.70 (0.06)	0.10 (0.03)	0.89 (0.03)	0.62 (0.10)	0.36 (0.07)
HEAVY-M	θ_1^M	θ_2^M	$\delta_{i,1}^M$	$\delta_{i,2}^M$	$\alpha_{i,1}^M$	$\alpha_{i,2}^M$
	0.49 (0.06)	0.50 (0.05)	0.15 (0.02)	0.84 (0.02)	0.50 (0.05)	0.49 (0.04)

Table 1: Full-sample parameter estimation results for the pair of the factor and COF. Robust standard errors are reported in parentheses.

Tables 2 presents the minimum, the 25th percentile, the mean, the median, the 75th percentile and the maximum of the parameter estimates across all firms. The estimates of the dynamics of the conditional β lie in a narrow range, with most sensitivity parameters estimated to be close to 0.10. The estimates of the parameters in the HEAVY-M equations fall into a particularly narrow band. The estimated dynamics of the conditional idiosyncratic volatilities are more dispersed, although mode series are highly responsive to idiosyncratic news. The estimates from the idiosyncratic component of the HEAVY-M are also more similar than the HEAVY-P, with both high sensitivity to news and large persistence.

In traditional ARCH-type models, both volatilities and β s are driven by daily shocks. We consider an augmented model of the HEAVY-P equations which allow for the natural daily shock to enter the model. The modified equations are

⁸In a small minority of series, either $\delta_{i,1}$ or $\alpha_{i,1}$ (or both) is not statistically significant. This may be due to constancy of conditional β and/or idiosyncratic volatility.

	β				Idiosyncratic			
	$\delta_{i,1}$	$\delta_{i,2}$	$\delta_{i,1}^M$	$\delta_{i,1}^M + \delta_{i,2}^M$	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,1}^M$	$\alpha_{i,1}^M + \alpha_{i,2}^M$
Min	0.01	0.01	0.07	0.96	0.07	0.06	0.08	0.90
Q _{.25}	0.08	0.84	0.12	0.98	0.42	0.29	0.39	0.99
Mean	0.12	0.82	0.15	0.98	0.51	0.44	0.44	0.99
Median	0.10	0.89	0.15	0.99	0.49	0.42	0.45	0.99
Q _{.75}	0.13	0.92	0.17	0.99	0.64	0.54	0.50	0.99
Max	0.53	0.99	0.26	0.99	0.92	0.93	0.63	1.00

Table 2: Cross-sectional statistics of full-sample parameter estimates for the 40 financial firms. The left panel contains estimates from the β component of the HEAVY-P and HEAVY-M equations, and the right panel contains estimates from the idiosyncratic volatility component.

$$\begin{aligned}
\sigma_{f,t}^2 &= \theta_0 + \theta_1 RM_{f,t-1} + \theta_2 \sigma_{f,t-1}^2 + \theta_3 r_{f,t-1}^2 \\
\beta_{i,t} &= \delta_{i,0} + \delta_{i,1} R\beta_{i,t-1} + \delta_{i,2} \beta_{i,t-1} + \delta_{i,3} r_{f,t-1} r_{i,t-1} / \sigma_{f,t-1}^2 \\
\sigma_{i,t}^2 &= \alpha_{i,0} + \alpha_{i,1} RIV_{i,t-1} + \alpha_{i,2} \sigma_{i,t-1}^2 + \alpha_{i,3} \epsilon_{i,t-1}^2
\end{aligned}$$

where θ_3 allows the daily factor return to influence the factor volatility, $\delta_{i,3}$ allows a daily β shock to enter the β dynamics and $\alpha_{i,3}$ allows for an idiosyncratic shock where $\epsilon_{i,t} = r_{i,t} - \beta_{i,t} r_{f,t}$. These were tested one-at-a-time, and all were conducted at the 5% level. Estimation was conducted using the 2-step estimator described in section 3.1, and inference was conducted using robust standard errors similar to those described in section 3.2. The daily shock in the factor volatility equation was not significant. In the β equations, $\delta_{i,3}$ was significant for 8 firms, but had a much smaller coefficient than $\delta_{i,1}$ indicating the changes in the fit value were typically small. In the idiosyncratic equations, $\alpha_{i,3}$ was significant for 7 firms.

We conducted a number of other experiments to assess whether all components of the models were necessary. We first compared the in-sample fit with that of the pairwise fit of the cDCC model, which has recently been used to fit dynamic β models (Engle & Kelly 2012, Bali, Engle & Tang 2012). The cDCC model specifies dynamics for the conditional variances of the market and the individual asset as standard GARCH models, and the conditional correlation is modeled using the standardized residuals.

The cDCC models the volatility of the market index and the firm's equity returns as standard GARCH(1,1) processes (Bollerslev 1986) so that

$$\begin{aligned}
\tilde{\sigma}_{f,t}^2 &= \omega_f^D + \alpha_f^D r_{f,t-1}^2 + \beta_f^D \tilde{\sigma}_{f,t-1}^2 \\
\tilde{\sigma}_{i,t}^2 &= \omega_i^D + \alpha_i^D r_{i,t-1}^2 + \beta_i^D \tilde{\sigma}_{i,t-1}^2
\end{aligned} \tag{23}$$

where $\tilde{\sigma}_{f,t}$ is used to distinguish the low-frequency factor volatility from the high-frequency-based factor volatility in the HEAVY model. The conditional correlation is modeled with the cDCC model (Aielli 2013), which is an extension of the DCC model (Engle 2002) which allows for consistent estimation in 3-steps. In cDCC models, the conditional covariance matrix of daily return vectors is decomposed as

$$\text{Var}_{t-1} \begin{pmatrix} r_{m,t} \\ r_{i,t} \end{pmatrix} = \begin{pmatrix} \sigma_{m,t} & 0 \\ 0 & \sigma_{i,t} \end{pmatrix} \begin{pmatrix} 1 & \rho_t \\ \rho_t & 1 \end{pmatrix} \begin{pmatrix} \sigma_{m,t} & 0 \\ 0 & \sigma_{i,t} \end{pmatrix}$$

The conditional correlation is modeled using a transformation of an auxiliary process,

$$\begin{aligned} Q_t &= (1 - \alpha_C - \beta_C)S + \alpha_C \tilde{\epsilon}_{t-1} \tilde{\epsilon}_{t-1}' + \beta_C Q_{t-1} \\ R_t &= (Q_t \odot I_2)^{-1/2} Q_t (Q_t \odot I_2)^{-1/2} \end{aligned}$$

where R_t denotes the correlation matrix; $\odot$ denotes Hadamard product and S is a symmetric matrix with diagonal elements equal to 1. $\tilde{\epsilon}_t = (Q_t \odot I_2)^{1/2} \epsilon_t$ is a “revolatilized” return vector with unit unconditional variance where $\epsilon_t = (\epsilon_{m,t}, \epsilon_{i,t})'$.

We also compared the factor HEAVY with the daily β -GARCH model (eq. (1)), as well as the multivariate HEAVY estimated on the market and each individual asset. Finally, we compared the factor HEAVY to two nested specifications. The first assumes that the conditional factor loading is constant, which corresponds to restrictions that $\delta_{i,1} = \delta_{i,2} = 0$, and the second which assumes that the idiosyncratic volatility is constant, corresponding to the restrictions $\alpha_{i,1} = \alpha_{i,2} = 0$.

Table 3 contains the value of the difference in the log-likelihood of these models,

$$\sum_{t=1}^T (l_{t,\text{Factor HEAVY}} - l_{t,\text{Alternative}})$$

evaluated using only daily data in the case of HEAVY models. Since all models assume that returns are conditionally normal, these are directly comparable, and positive values indicate that the factor HEAVY produced a superior in-sample fit. The Factor HEAVY produces much larger log-likelihoods than either of the daily-only models, and the closest daily model differs by over 80 log-likelihood points. Moreover, the typical difference is more than 150 points, indicating that there is substantial information in the high-frequency measures. The models nested in the factor HEAVY – either with a constant factor loading or a constant idiosyncratic volatility – are also considerable worse, although the assumption of constant factor loading does not always lead to large changes in the log-likelihood. On the other hand, idiosyncratic volatilities appear to be time-varying with typical log-likelihood differences of 1,000 points. Finally, the results comparing the Factor HEAVY with bivariate multivariate HEAVY models indicate some preference for the factor HEAVY, with better performance in more than 75% of the series examined.

We also examined the fit of the model using residual-based tests. We are primarily concerned with misspecification of the conditional covariance between the returns on a pair of equities since the model does not explicitly attempt to fit these. We are interested in testing the null that

$$H_0 : E_{t-1} (r_{i,t} r_{j,t}) = \beta_{i,t} \beta_{j,t} \sigma_{f,t}^2,$$

	Daily		High-Frequency Based		
	cDCC	β GARCH	Constant Factor Loading	Constant Idiosyncratic Variance	Multivariate HEAVY
Min	85.5	82.48	4.277	211.6	-23.340
25%	137.9	135.8	31.77	483.9	1.220
Median	157.4	164.6	65.39	956.6	6.000
Mean	169.6	169.8	67.33	1004	6.356
75%	201.0	190.7	101.7	1206	9.074
Max	307.0	309.7	149.0	2441	42.05

Table 3: In-sample log-likelihood comparison between Factor HEAVY and competing models for all equities. All comparisons were performed using the pair of the market and one of the financial firms.

where we have suppressed the dependence of the factor loadings and market variance on the model parameters,

$$\zeta = (\theta'_o, \delta'_{i,o}, \delta'_{j,o})'$$

which includes the parameters of the conditional variance of the factor as well as the conditional factor loadings of assets i and j . Since the parameter estimates are variation-free across models, it is not necessary to include the parameters for the other components in ζ .

We use the robust regression-based specification tests developed in [Wooldridge \(1990\)](#). We use $z_{i,j,t} = (1, r_{i,t-1}r_{j,t-1}, r_{i,t-2}r_{j,t-2})$ as the vector of misspecification indicator variables. These indicators will have power against static misspecification (through the constant term) as well as persistence in the cross-products. The test is implemented in two steps. First, we regress each element of $z_{i,j,t}$ on the gradient vector $\nabla_{\zeta}\eta_t$ of $\eta_t = r_{i,t}r_{j,t} - \beta_{i,t}\beta_{j,t}\sigma_{f,t}^2$ with respect to ζ evaluated at the estimate $\hat{\zeta}$. We define $\hat{z}_{i,j,t}$ as the residual vector obtained from this regression. Second, we regress unit on the vector $\hat{\eta}_t\hat{z}_{i,j,t}$ where $\hat{\eta}_t$ is the value of η_t evaluated at $\hat{\zeta}$. The test statistic is then computed as $T \times R^2$ where the R^2 comes from the second regression. The test statistics is asymptotically distributed as χ_3^2 . Figure 3 contains a histogram of the test statistics. We fail to reject the null of correct conditional specification in 93% of the pairs tested (the 5% critical value of a χ_3^2 is 7.81), and so the rejection rate is close to size.

4.3 Out-of-Sample Forecasting Results

We next turn attention to assessing the out-of-sample performance of the factor HEAVY, and we focus the out-of-sample comparisons on the cDCC model, as a leading example of low-frequency models. All models are fitted using a recursive scheme where parameters are updated once a

Figure 3: Histogram of misspecification test statistics.

week (on Fridays) starting from July 2005, which allows for a minimum of 1,250 days in the smallest models. The cDCC models were always estimated pairwise with only the market and one of the financial firms. We will evaluate the performance of the model using 1-day, 1-week and 2-week forecast horizons, which are all important for risk management.⁹

Figure 4 shows the 1-step forecasts of cDCC and factor HEAVY models for COF. The conditional β forecasts from the factor HEAVY are more persistent than those from the cDCC. This higher persistence originates from the direct modeling of conditional β in eq. (7) where the coefficients indicate considerable persistence and a slow response to news. Both series of conditional β s show large increases during the financial crisis and the continuing turmoil of early 2009, although they differ markedly in the pre-crisis period of 2007 until mid-2008, where the HEAVY model indicates substantial increases in conditional β while the cDCC does not.

4.3.1 Statistical Accuracy

We compare the statistical accuracy of all models using out-of-sample comparisons. The QLIK loss function has recently emerged as a sensible method to evaluate variance and covariance forecasts in the presence of noisy proxies (see [Patton & Sheppard \(2009\)](#), [Patton \(2011\)](#) and [Laurent, Rombouts & Violante \(2009\)](#)). The QLIK loss function uses the kernel of the Gaussian log-likelihood to evaluate forecasts,

$$l_t^s(\hat{\Sigma}_{t+s|t}) = \ln |\hat{\Sigma}_{t+s|t}| + \hat{\Sigma}_{t+s|t}^{-1} \text{tr}(C_{t+s}). \quad (24)$$

⁹Results for longer horizons, up to 1 months, indicate there is little difference between the models.

Figure 4: 1-step β forecasts of cDCC and factor HEAVY models for COE. Solid lines denote the factor HEAVY forecasts; while dash lines represent the cDCC forecasts.

where s is the number of steps ahead and $\hat{\Sigma}_{t+s|t}$ is a model-based time- t conditional forecast of the covariance of r_{t+s} and C_{t+s} is the proxy for the unobserved covariance at time $t + s$. The natural proxy is $r_t r'_t$, although others, such as a realized measure, can be used.¹⁰ Suppressing the explicit dependence on $\hat{\Sigma}_{t+s|t}$, the QLIK loss function can be decomposed to isolate various component losses,

$$\begin{aligned}
l_t^s &= \underbrace{\left(\ln(\sigma_{f,t+s|t}^2) - \frac{V_{f,t+s}}{\sigma_{f,t+s|t}^2} \right)}_{\text{factor}} + \sum_{i=1}^N \underbrace{\left(\ln(\sigma_{i,t+s|t}^2) - \frac{\beta_{i,t+s|t}^2 V_{f,t+s} - 2\beta_{i,t+s|t} C_{fi,t+s} + V_{i,t+s}}{\sigma_{i,t+s|t}^2} \right)}_{\text{idiosyncratic } i} \quad (25) \\
&= \underbrace{\left(\ln(\sigma_{f,t+s|t}^2) - \frac{V_{f,t+s}}{\sigma_{f,t+s|t}^2} \right)}_{\text{factor}} + \sum_{i=1}^N \underbrace{\left(\ln(\beta_{i,t+s|t}^2 \sigma_{f,t+s|t}^2 + \sigma_{i,t+s|t}^2) - \frac{V_{i,t+s}}{\beta_{i,t+s|t}^2 \sigma_{f,t+s|t}^2 + \sigma_{i,t+s|t}^2} \right)}_{\text{individual equity } i} \\
&\quad + \text{term corresponding to the dependence structure} \\
&= l_{f,t}^s + \sum_{i=1}^N l_{i,t}^s + l_{C,t}^s
\end{aligned}$$

where $V_{f,t+s}$, $V_{i,t+s}$ and $C_{fi,t+s}$ are the proxies for the factor variance, equity variance and their covariance on day $t + s$. While this decomposition is generally only applicable to factor models, it also holds for any bivariate model which contains the factor, such as the daily cDCC models estimated.

All comparisons are implemented using [Diebold & Mariano \(1995\)](#) (DM) tests of equal predictive accuracy where the covariance forecasts come from the factor HEAVY and bivariate cDCC models. The null in DM test statistics is of equal expected loss,

$$H_0 : E[l_t^s(\hat{\Sigma}_{t+s|t}^{\text{HEAVY}})] = E[l_t^s(\hat{\Sigma}_{t+s|t}^{\text{cDCC}})]$$

and the alternative is two-sided. Define the loss differential as $\delta_t = l_t^s(\hat{\Sigma}_{t+s|t}^{\text{cDCC}}) - l_t^s(\hat{\Sigma}_{t+s|t}^{\text{HEAVY}})$, then positive values correspond to superior performance of the factor HEAVY while negative values correspond to superior performance of the cDCC. The test statistic is implemented as a simple t-statistic

$$DM = \sqrt{T - \tau + 1} \frac{\sum_{t=\tau}^T \delta_t}{\sqrt{\text{avar}(\delta_t)}}$$

where $\text{avar}(\delta_t)$ is a consistent estimator of the long-run variance of δ_t . This is estimated using a Newey-West covariance estimator with $\min(10, \lfloor 1.2T^{1/3} \rfloor)$ lags.

¹⁰The proxy should satisfy the condition that $E_t(C_{t+s}) = \Sigma_{t+s}$, so that it is unbiased. When forecasting the covariance over the entire day, most realized measures will fail to meet this criteria.

Table 4 contains the value of the DM test statistics for three distinct comparisons. The first compares the bivariate loss (eq. (24)) applied to the market and the individual firm, the firm only loss ($l_{i,t}^s$) and the copula ($l_{C,t}^s$). All comparisons were conducted for 1-day, 5-day and 2-week horizons, and are implemented on point-wise forecasts and not cumulative – that is, predicting $\Sigma_{t+s|t}$ not $\sum_{j=1}^s \Sigma_{t+j|t}$. The out-performance of these models is so strong at short horizons that cumulative comparisons are not useful. At the 1-day horizon, the factor HEAVY model is dominant in both the joint and firm tests, with strong rejections for almost all pairs. The copula based-comparisons are more mixed, although the null is rejected in favor of the factor HEAVY for 12 of the 40 series, and only once in favor of cDCC. The results using 1-week forecasts are unsurprisingly weaker, although the factor HEAVY is still preferred. Finally at the 2-week horizon, a similar pattern is found, with many statistically significant rejections both in the joint and firm volatility. The DM tests for the factor variance also uniformly preferred the HEAVY model, and the test statistic at the 1-day horizon was 3.72.

The ability of models to forecast conditional β was also examined using the estimated idiosyncratic variance as the loss function. The conditional β represents the optimal hedge ratio and so an accurate forecast should lead to a small tracking error. This leads to the loss function

$$l_{i,t}^s = (r_{i,t+s} - \beta_{i,t+s|t} r_{f,t+s})^2.$$

Table 5 contains the result of DM tests of the null of equal idiosyncratic variance of the residual, comparing the factor HEAVY to the cDCC. In most cases, the null cannot be rejected. However, when it is, the rejections indicate superior performance of the factor HEAVY. These results are weaker than in the QLIK tests, which may be due to heteroskedasticity of the idiosyncratic error. The statistical tests all indicate that the out-of-sample performance of the factor HEAVY is superior to the leading low-frequency model. Moreover, the gains are particularly striking in terms of predicting variance. The gains in terms of predicting dependence, whether from the copula term which arises from the QLIK loss function, or from measuring the idiosyncratic variance, are smaller. We hypothesize that the gains to forecasting volatility may be larger since the range of conditional variances is much larger than that of conditional correlations or β s, which should aid in finding the superior model.

4.3.2 Marginal Expected Shortfall Forecasting

Risk forecasting, and in particular systemic risk forecasting, is an important application of multivariate volatility models. Marginal expected shortfall, introduced by [Acharya et al. \(2012\)](#), measures the expected loss conditional on a factor being in a state of stress, and is defined

	1 day			1 week			2 weeks		
	Joint	Firm	Copula	Joint	Firm	Copula	Joint	Firm	Copula
AET	4.42	3.25	1.44	2.40	2.64	-0.39	2.82	2.39	-0.53
AFL	3.14	1.79	3.30	1.37	1.35	0.36	1.62	1.47	-0.05
AIG	-0.09	-1.00	1.20	-0.81	-1.27	1.60	-1.17	-1.42	-0.07
ALL	3.21	2.25	0.98	1.99	1.61	1.45	2.39	1.90	1.09
AXP	3.88	2.18	1.48	1.55	0.99	1.42	1.04	0.42	0.69
BAC	1.37	0.81	-0.24	-0.27	-0.67	0.41	-0.70	-1.07	0.34
BBT	4.22	2.61	1.06	1.05	0.36	0.99	2.45	2.33	0.84
BEN	3.70	2.10	0.72	1.86	1.24	1.14	2.11	1.43	1.08
BK	5.14	2.65	2.77	2.15	1.91	0.58	2.28	2.05	0.72
C	2.45	1.02	1.41	0.28	-0.73	1.37	-1.11	-1.77	-0.19
CB	5.56	2.96	4.69	2.73	1.63	1.89	3.40	1.99	0.47
CMA	4.00	2.06	1.86	2.36	1.64	1.68	3.16	2.91	1.45
COF	5.90	3.70	2.35	2.93	2.86	0.95	1.83	2.24	0.14
ETFC	1.88	2.08	0.01	1.36	1.66	0.76	1.81	1.65	1.22
EV	2.87	0.83	0.34	0.79	-0.70	1.61	1.69	0.97	0.47
FITB	4.07	2.69	2.01	0.98	0.29	1.58	1.86	1.28	1.03
FNM	1.37	1.12	0.44	1.38	1.11	1.31	1.70	1.46	0.55
FRE	0.88	0.99	0.00	-0.38	-0.59	1.36	-0.49	-0.73	1.50
GS	1.44	1.51	-1.30	0.65	1.76	-1.09	0.15	0.12	-0.48
HIG	0.85	1.99	-1.21	-0.41	-0.15	-0.73	1.13	1.30	0.59
HRB	2.73	2.34	-0.08	0.70	0.65	-0.04	-0.27	-0.20	-0.91
JNS	4.02	2.27	2.64	1.32	1.12	0.73	1.02	0.91	0.81
JPM	4.74	2.81	1.39	0.77	0.48	0.76	0.15	-0.36	0.66
KEY	2.24	-0.05	1.88	-0.29	-1.23	0.85	0.02	-1.05	1.03
LEH	2.05	1.94	0.17	1.20	1.87	0.29	-0.93	-0.79	-1.72
LM	3.08	1.28	0.92	2.04	2.28	-0.10	3.16	2.54	0.29
MER	4.34	3.99	1.75	1.08	0.83	0.76	-1.21	-1.16	-1.40
MET	5.00	1.72	3.10	2.15	1.10	1.99	0.92	0.66	0.30
MMC	5.34	3.63	0.25	0.61	3.93	-1.16	0.95	3.05	-1.12
MS	3.01	2.60	-0.61	0.47	0.94	-0.62	-0.98	-0.87	-1.30
PGR	2.17	2.43	-1.66	0.35	1.75	-1.61	0.52	1.49	-2.14
PNC	4.74	2.82	2.31	1.82	0.47	1.46	0.36	-0.13	1.85
PRU	3.13	2.71	-1.07	-0.46	0.57	-1.91	0.49	0.32	-0.19
SCHW	3.22	2.55	-0.54	1.36	1.52	0.11	0.89	0.64	0.00
SLM	0.01	-0.51	0.53	-0.16	-0.24	-0.02	0.91	1.42	-1.28
STI	3.56	2.64	0.69	0.78	0.48	0.49	0.11	0.00	-0.23
STT	3.78	2.12	3.14	0.22	0.70	-0.27	0.70	0.93	0.00
TROW	2.53	0.44	0.62	-0.09	-0.80	0.45	0.98	-0.08	0.97
UNH	2.76	1.23	0.88	1.24	0.77	1.14	1.41	0.86	0.49
USB	2.01	0.69	0.10	-0.27	-0.70	0.22	-0.59	-0.71	-0.39

Table 4: Diebold-Mariano statistics for the null hypothesis that factor HEAVY and cDCC forecasts are equally accurate when forecasting 1-day, 1-week and 2-week ahead volatility of firm equity returns, the dependence structure and the covariance matrix. Numbers in bold font indicate that factor HEAVY forecasts are more accurate than cDCC forecasts at the 10% significance level. Italic numbers indicate the converse.

	1 day	1 week	2 weeks		1 day	1 week	2 weeks
AET	1.56	1.71	1.83	HRB	1.75	1.79	1.73
AFL	2.10	2.33	2.70	JNS	1.85	1.76	0.89
AIG	1.07	1.24	1.14	JPM	-0.27	1.09	1.17
ALL	-0.07	0.73	0.76	KEY	0.64	1.04	1.16
AXP	-0.01	0.37	0.96	LEH	-1.08	-0.91	-0.96
BAC	-0.12	0.30	0.57	LM	1.92	1.92	2.23
BBT	1.12	1.20	1.36	MER	1.66	1.70	1.23
BEN	1.35	2.04	2.18	MET	0.60	0.66	1.24
BK	0.77	2.81	2.68	MMC	1.07	0.92	1.13
C	0.79	0.80	0.83	MS	1.35	1.13	0.61
CB	0.01	1.71	2.27	PGR	1.00	1.10	1.22
CMA	1.06	1.65	1.99	PNC	-1.27	-0.97	0.40
COF	1.70	2.31	2.60	PRU	-0.91	-0.62	-0.65
ETFC	-0.73	0.17	1.61	SCHW	-0.65	-0.61	-0.78
EV	-0.18	0.62	1.28	SLM	1.98	2.54	2.25
FITB	-0.35	0.72	1.48	STI	0.87	1.55	1.83
FNM	-0.76	-0.72	-0.29	STT	1.09	1.33	2.15
FRE	-1.12	-0.14	1.06	TROW	-0.91	-0.59	0.91
GS	0.82	0.06	0.10	UNH	1.11	0.88	1.33
HIG	1.30	0.85	1.17	USB	-0.99	-0.10	0.99

Table 5: Diebold-Mariano statistics for the null hypothesis that factor HEAVY and cDCC forecasts are equally accurate when forecasting 1-day, 1-week and 2-week ahead β s. Numbers in bold font indicate that factor HEAVY forecasts are more accurate than cDCC forecasts at the 10% significance level. Italic numbers indicate the converse.

$$\text{MES}_{i,t} = -E_t(r_{i,t+1}|r_{f,t+1} < c)$$

where c is a threshold which determine whether the factor is showing signs of stress. The longer term s -step ahead MES is defined similarly but based on cumulative returns:

$$\text{MES}_{i,t}^s = -E_t(R_{i,t+1:s}|R_{f,t+1:s} < c_s)$$

where $R_{i,t+1:s} = \exp(\sum_{\tau=1}^s r_{i,t+\tau}) - 1$ is the cumulative return on firm equity from $t + 1$ to $t + s$, $R_{f,t+1:s}$ is similarly defined and c_s allows for explicit dependence between the horizon and the threshold for a stress event.

We follow a similar procedure to that introduced in [Brownlees & Engle \(2012\)](#) to estimate MES. The 1-step MES is particularly simple to compute, and is

$$\begin{aligned}\text{MES}_{i,t} &= -\mathbb{E}_t \left(\beta_{i,t+1} r_{f,t+1} + \epsilon_{i,t+1} \mid r_{f,t+1} < c \right) \\ &= -\beta_{i,t+1} \sigma_{f,t+1} \mathbb{E}_t \left(\xi_{f,t+1} \mid \xi_{f,t+1} < c/\sigma_{f,t+1} \right) - \sigma_{i,t+1} \mathbb{E}_t \left(\xi_{i,t+1} \mid \xi_{f,t+1} < c/\sigma_{f,t+1} \right).\end{aligned}$$

In the cDCC model, $\beta_{i,t} = \tilde{\sigma}_{i,t} \rho_t / \tilde{\sigma}_{f,t}$, $\sigma_{i,t} = \tilde{\sigma}_{i,t} \sqrt{1 - \rho_t^2}$ and $\xi_{i,t} = (r_{i,t} - \beta_{i,t} r_{f,t}) / \sigma_{i,t}$. $\mathbb{E}_t \left(\xi_{f,t+1} \mid \xi_{f,t+1} < c/\sigma_{f,t+1} \right)$ and $\mathbb{E}_t \left(\xi_{i,t+1} \mid \xi_{f,t+1} < c/\sigma_{f,t+1} \right)$ are estimated using a nonparametric kernel estimation similar to that in [Brownlees & Engle \(2012\)](#),

$$\hat{\mathbb{E}}_t \left(x_{t+1} \mid \xi_{f,t+1} < c/\sigma_{f,t+1} \right) = \frac{\sum_{\tau=1}^t K_l \left(\frac{c}{\sigma_{f,t+1}} - \xi_{f,\tau} \right) x_\tau}{\sum_{\tau=1}^t K_l \left(\frac{C}{\sigma_{f,t+1}} - \xi_{f,\tau} \right)} \quad (26)$$

where x is either ξ_f or ξ_i , $K_l(z) = \int_{-\infty}^{z/l} k(u) du$, $k(\cdot)$ is the Gaussian kernel and l is the bandwidth.

Longer-horizon MES forecasting is more complex and requires the use of a bootstrap to simulate returns. The return simulation is constructed using standardized residuals, and, in the factor HEAVY, standardized realized measures. The s -step MES requires an estimate of the expected (cumulative) loss conditional on the factor suffering a large (cumulative) loss. In the factor HEAVY, we construct the tuple of standardized innovations

$$\boldsymbol{\nu}_t = (\xi_{f,t}, \xi_{i,t}, \text{vech}(\Xi_{i,t}))'$$

where

$$\Xi_t = M_{i,t}^{-1/2} (R M_{i,t}) M_{i,t}^{-1/2}.$$

The values of the tuple $\boldsymbol{\nu}_t$ are then used to simulate from the factor HEAVY for a longer horizon, and simulated cumulative returns are constructed for both the factor and the asset. The s -step MES is then constructed as

$$\text{MES}_{i,t}^s = - \frac{\sum_{b=1}^B \hat{R}_{i,t+1:s}^b I \left[\hat{R}_{f,t+1:s}^b < c_s \right]}{\sum_{b=1}^B I \left[\hat{R}_{f,t+1:s}^b < c_s \right]} \quad (27)$$

where $I[x] = 1$ if condition x is satisfied and equal to zero otherwise and $\hat{R}_{i,t+1:s}$ is used to denote

that these are simulated returns. This is repeated across all assets using only pair-wise measures. In the cDCC, the standardized innovations include only $(\xi_{f,t}, \xi_{i,t})$ where the standardization is computed using the cDCC-fit values of the conditional covariance. The simulation procedure is otherwise identical. The entire procedure is recursively implemented so that all estimates including both the parameters of the conditional covariance models as well as the other inputs to the MES are “real-time”.

MES prediction accuracy is evaluated using Relative Mean Square Error (RMSE), which is used in place of MSE to control for strong heteroskedasticity of time-varying losses. The RMSE for s -step ahead MES is defined as

$$\text{RMSE}_{i,t_c}^s = \left(\frac{-R_{i,t_c+1:s} - \text{MES}_{i,t_c}^s}{\text{MES}_{i,t_c}^s} \right)^2$$

where the period from $t_c + 1$ to $t_c + s$ represents the event period of length s during which the market experiences a loss larger than a pre-determined threshold. Since MES can only be evaluated when the market suffers a relatively large loss, few data points are available to evaluate the models.

Relative predictive accuracy is assessed using Diebold-Mariano tests where the loss differential is $\delta_{i,t_c} = \text{RMSE}_{i,t_c}^{cDCC} - \text{RMSE}_{i,t_c}^{HEAVY}$. The test statistic is

$$\sqrt{T_c} \frac{\bar{\delta}}{\sqrt{\text{avar}(\delta_t)}}$$

where T_c denotes the number of event days.

Figure 5 shows the 1-step MES forecasts for COF from both the factor HEAVY and the cDCC, along with the market return, on days which are events in the out-of-sample period. For each event day, the 1-day, 1-week and 2-week return are shown. A large number of events occur during the final 4 months of 2008, which is highlighted.

Table 6 presents the DM statistics of equal predictive ability for each firm’s return. Factor HEAVY forecasts are more accurate on average than cDCC forecasts for all horizons we consider. The advantage is larger for longer horizons, which differs from the results in the volatility forecasting section. This difference arises since medium-term MES forecasts are cumulative and so superior performance of HEAVY models at shorter horizons accumulates.

Figure 5: Event days and the beginning days of event periods from July 2005 and the corresponding market index losses. Event days are the days on which the daily market index return declines by at least 2%. Event periods are those during which the market index experiences more than the corresponding threshold loss. Actual equity losses of Capital One (COF) and 1-day MES forecasts on event days from July 2005. Solid lines denote the actual losses; circle and star markers represent factor HEAVY and cDCC forecasts, respectively.

Figure 5 shows the actual event days and event periods and the corresponding market index losses. The corresponding threshold market index losses are 2%, 4% and 6.5%, respectively.¹¹ Under these settings, there are 84 event days, 90 1-week event periods and 51 2-week event periods in the sample. Figure 5 compares MES forecasts with the losses for COF on event days.¹²

¹¹These values were chosen since larger values of the threshold index require extremely large bootstrap samples, especially in the period when volatility is low.

¹²The bandwidth in kernel estimation is set equal to 0.05. This value is chosen to give stable results.

While it is not clear from the figure which model performs better on average, factor HEAVY forecasts are more accurate at the 10% significance level. The figure shows that each model can track actual losses well on some event days. At the crisis peak around autumn 2008, however, factor HEAVY forecasts outperform cDCC forecasts where cDCC forecasts tend to seriously underestimate the actual losses. Forecasts from factor HEAVY models respond quickly to the shock while forecasts from cDCC models adjust slowly. This fact may provide insight into the source of the forecasting gains of factor HEAVY models.

Table 6 presents the Diebold-Mariano statistics associated with the 1-day, 1-week and 2-week MES forecasts.¹³ 1-day factor HEAVY forecasts are more accurate than the cDCC forecasts for 23 firms, and 10 of these are significant at the 10% level. On the other hand, cDCC forecasts are significantly more accurate for 4 firms. When looking at the 1-week MES, factor HEAVY still enjoys an advantage over cDCC in terms of the number of significant forecasts, but the advantage is somewhat smaller than for 1-day-ahead forecasts. The gains from using factor HEAVY models, however, are further reduced at the 2-week horizon, although for some firms like COF, factor HEAVY forecasts are significantly more accurate than cDCC forecasts at all forecast horizons of interests.

¹³When doing bootstrapping for s -step forecasting, we fix the number of remaining simulation paths to 1000 which satisfy $R_{f,t+1:s} < c_s$. Under this setting, the MES simulation result was found to be stable.

	1 day	1 week	2 weeks		1 day	1 week	2 weeks
AET	3.80	2.34	2.92	HRB	0.04	1.98	2.74
AFL	2.83	0.81	1.41	JNS	2.58	2.02	2.00
AIG	-3.17	-1.48	0.31	JPM	0.89	0.29	-1.01
ALL	0.20	-0.73	-1.84	KEY	-2.22	-1.58	-2.38
AXP	0.56	-0.99	-1.83	LEH	-0.96	-1.21	-1.39
BAC	-2.82	-1.72	-1.75	LM	2.71	2.08	1.19
BBT	1.31	-0.51	0.78	MER	-1.51	-1.19	-0.44
BEN	0.79	-1.14	-0.70	MET	-1.28	-1.19	-2.05
BK	2.72	0.73	0.04	MMC	2.95	2.70	2.82
C	-1.60	-1.87	-2.44	MS	0.65	-1.07	-2.32
CB	1.52	1.33	0.18	PGR	0.89	0.76	1.10
CMA	1.95	0.94	1.54	PNC	0.13	-1.16	-1.01
COF	3.16	2.19	2.26	PRU	-1.34	-1.64	-2.17
ETFC	-0.43	-0.21	-1.61	SCHW	3.16	1.95	2.27
EV	-0.67	-1.09	-1.62	SLM	1.00	-1.03	3.36
FITB	-1.61	-1.72	-1.94	STI	-1.57	-1.98	-2.36
FNM	-1.36	-0.92	1.22	STT	1.16	0.24	-1.05
FRE	-2.98	0.36	-1.12	TROW	-0.35	-1.00	-0.66
GS	0.98	0.71	0.46	UNH	1.66	2.81	2.47
HIG	-1.18	-1.10	0.99	USB	-0.14	-2.05	-1.48

Table 6: Diebold-Mariano statistics associated with the 1-day, 1-week and 2-week MES forecasts. Numbers in bold font indicate that factor HEAVY forecasts are more accurate than cDCC forecasts at the 10% significance level for this firm. Italic numbers indicate the superior performance by cDCC forecasts.

It is noted that COF has consistently good performance when forecasting volatility, covolatility, β and MES. Since MES is the expected equity loss of the firm conditional on a threshold market index loss, we suspect that these all play a role in the 1-day MES gains, although the performance in MES forecasting is stronger in models where the factor HEAVY performed better in terms of the copula. Factor HEAVY models are significantly more accurate when forecasting both the MES and copula for the firms such as AET, AFL, CMA, COF, JNS, LM, MMC and UNH. This relation between MES and copula forecasting remains but becomes weakens at longer horizons.

4.4 Extensions

Three extensions to the one-factor HEAVY model were explored. The first examines the gains to including an additional factor, where an industry factor, constructed as a portfolio of the included firms, is added to the model. The second examines the effect of changing the realized measure. The third examines the evidence for including asymmetries in the model.

4.4.1 Two factors

The one-factor specification is augmented to a 2-factor structure where the second factor is a constructed industry portfolio. The portfolio weights were computed using principle component analysis on CAPM residuals on the daily data. Define $\hat{\epsilon}_{i,t} = r_{i,t} - \hat{\beta}_i r_{f,t}$ to be CAPM residuals where $\hat{\beta}_i$ is the usual OLS estimator. A portfolio was then formed using the weight vector w which corresponds to eigenvector associated with the largest eigenvalue in the outer-product of $\hat{\Omega} = T^{-1} \sum_{t=1}^T \hat{\epsilon}_t \hat{\epsilon}_t'$. The new factor has daily returns $r_{PCA,t} = \sum w_i r_{i,t}$. High-frequency returns were similarly constructed using the same weight vector applied to the intra-daily data. Finally, the 10-minute realized covariances were computed using the high frequency returns of the constructed portfolio, along with the market and each individual firm.

The two-factor model was estimated using SPY and the PCA-factor for COF. Table 7 shows that all estimates are statistically significant except for $\delta_{i,1,1}^M$ and $\delta_{i,1,2}^M$. The estimates in the β and idiosyncratic volatility equations are similar to those in the one-factor model. The fit was also examined for all assets based on either the marginal log-likelihood of the firm or the bivariate log-likelihood of the market and the firm. We found that while model with 2 factors improved univariate the fit for 29 out of the 40 firms, the joint likelihood including the firm and the market only improved in 13 cases. We interpret these findings as weak evidence that a second factor is not required for this data set – less homogeneous panels, such as those that include firms across sectors or with other important differences in characteristics, would be more likely to require multiple factors.

	Factor				β				Idiosyncratic	
HEAVY-P	A_{11}	A_{22}	B_{11}	B_{22}	$\delta_{i,1,1}$	$\delta_{i,2,1}$	$\delta_{i,1,2}$	$\delta_{i,2,2}$	$\alpha_{i,1}$	$\alpha_{i,2}$
	0.47 (0.18)	0.41 (0.09)	0.87 (0.12)	0.91 (0.05)	0.11 (0.04)	0.84 (0.06)	0.18 (0.06)	0.80 (0.10)	0.65 (0.09)	0.34 (0.08)
HEAVY-M	A_{11}^M	A_{22}^M	B_{11}^M	B_{22}^M	$\delta_{i,1,1}^M$	$\delta_{i,2,1}^M$	$\delta_{i,1,2}^M$	$\delta_{i,2,2}^M$	$\alpha_{i,1}^M$	$\alpha_{i,2}^M$
	0.59 (0.05)	0.51 (0.02)	0.79 (0.05)	0.86 (0.01)	0.11 (0.15)	0.88 (0.24)	0.10 (0.18)	0.88 (0.30)	0.50 (0.04)	0.49 (0.05)

Table 7: Parameter estimation results of a two-factor factor HEAVY model for COF, over the whole sample period. Robust standard errors are reported in parentheses.

4.4.2 Alternative Realized Measures

As a robustness check, the model was fit using alternative realized measures. The top panel of table 8 report parameter estimates and the (joint) log-likelihood across 5 choices of realized measure: realized covariance sampled between 5 and 30 minutes, all using 1-minute subsampling,

and multivariate realized kernels. The second panel reports and DM test statistics from 1-day forecasts for COF against the cDCC model using the same measures.. All results are stable across alternative realized measures, and realized measures which sample more often perform somewhat better both in sample and out-of sample, with realized kernel producing the best in-sample fit. The key parameters which measure the sensitivity to the realized measure in the HEAVY-P equations, θ_1 (market) , $\delta_{i,1}$ (factor loading) and $\alpha_{i,1}$ (idiosyncratic variance), all decline as the realized measure deteriorates, which is consistent with additional measurement noise when sampling less-frequently. All estimators consistently out-perform cDCC forecasts, and a similar pattern appears in terms of using more precise estimators.

The bottom three panels of table 8 show the relative distribution of parameters using the 25th quantile, median and 75th quantile of the estimates across all 40 assets, as well as the distribution of likelihood differences relative to the likelihood of the 10-minute realized covariance-based model (negative indicates better performance of the 10-minute model). These are broadly consistent with the pattern for COF, where less frequent sampling leads to smaller sensitivity to news as well as smaller likelihoods. The performance of the 5-minute and 10-minute realized covariance was extremely similar.

	In-Sample				
	5-minute	10-minute	15-minute	30-minute	Kernel
θ_1	0.28	0.26	0.26	0.24	0.28
$\delta_{i,1}$	0.12	0.10	0.10	0.09	0.13
$\alpha_{i,1}$	0.64	0.62	0.60	0.52	0.63
Likelihood	-9125	-9138	-9163	-9229	-9115

	Out-of-sample				
	5-minute	10-minute	15-minute	30-minute	Kernel
Joint	5.93	5.90	6.07	6.56	6.03
Factor	3.57	3.72	3.74	3.99	3.74
Firm	3.67	3.70	3.75	3.89	3.77
Copula	2.36	2.35	2.54	3.09	2.42
β	1.81	1.70	1.54	1.19	1.78
MES	2.89	3.16	3.12	2.96	2.71

5-miunte realized covariance							
	$\delta_{i,1}$	$\delta_{i,2}$	$\delta_{i,1}^M$	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,1}^M$	ΔL
$Q_{.25}$	0.07	0.74	0.15	0.41	0.25	0.41	-2.94
Median	0.11	0.86	0.18	0.54	0.38	0.48	-1.00
$Q_{.75}$	0.17	0.91	0.21	0.69	0.48	0.52	3.78

15-minute realized covariance							
	$\delta_{i,1}$	$\delta_{i,2}$	$\delta_{i,1}^M$	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,1}^M$	ΔL
$Q_{.25}$	0.06	0.85	0.10	0.35	0.33	0.35	-4.27
Median	0.08	0.90	0.12	0.45	0.49	0.41	-1.99
$Q_{.75}$	0.11	0.92	0.14	0.59	0.59	0.47	1.46

30-minute realized covariance							
	$\delta_{i,1}$	$\delta_{i,2}$	$\delta_{i,1}^M$	$\alpha_{i,1}$	$\alpha_{i,2}$	$\alpha_{i,1}^M$	ΔL
$Q_{.25}$	0.05	0.90	0.06	0.24	0.46	0.29	-19.33
Median	0.06	0.92	0.08	0.34	0.61	0.34	-14.79
$Q_{.75}$	0.08	0.94	0.10	0.53	0.71	0.40	-5.97

Table 8: Parameter estimates, in-sample likelihood and Diebold-Mariano statistics of 1-day forecasts for COF using different realized measures, i.e. 5-minute, 10-minute, 15-minute and 30-minute realized covariance matrices with 1-minute subsampling, as well as realized kernels.

4.4.3 Conditional Asymmetries in Volatility

The third extension is to examine whether asymmetries are needed in the model. GJR-GARCH and TARARCH models have been broadly found to fit conditional variances better than symmetric

GARCH models (Glosten et al. 1993, Zakoian 1994). A GJR-like specification can be constructed in the HEAVY-P using an indicator variable based on the daily return,

$$\sigma_{f,t}^2 = \theta_0 + \theta_1 RM_{f,t-1} + \theta_2 \sigma_{f,t-1}^2 + \theta_3 RM_{f,t-1} I(r_{f,t-1} < 0) \quad (28)$$

We tested the null that $\theta_3 = 0$ which was strongly rejected. Without the asymmetry (imposing $\theta_3 = 0$), $\hat{\theta}_1 = 0.26$. When the asymmetry is introduced, $\hat{\theta}_1 = 0.12$ and $\hat{\theta}_3 = 0.19$. These coefficient changes are large, and the full-sample marginal log-likelihood increased from -3674 to -3611.

Following Braun et al. (1995b), asymmetries were added to the conditional factor loading through three terms,

$$\begin{aligned} \beta_{i,t} = & \delta_{i,0} + \delta_{i,1} R\beta_{i,t-1} + \delta_{i,2} \beta_{i,t-1} + \delta_{i,3} R\beta_{i,t-1} I(r_{f,t-1} < 0) \\ & + \delta_{i,4} R\beta_{i,t-1} I(\epsilon_{i,t-1} < 0) + \delta_{i,5} R\beta_{i,t-1} I(r_{f,t-1} \epsilon_{i,t-1} < 0) \end{aligned}$$

where $\delta_{i,3}$ allows for asymmetries due to the market return, $\delta_{i,4}$ allows asymmetries through the idiosyncratic shock and $\delta_{i,5}$ allows asymmetries through the interaction of the market return and the idiosyncratic shock. This specification was fit to all 40 firms and the null $\delta_{i,3} = \delta_{i,4} = \delta_{i,5} = 0$ was tested using a Wald test with a sandwich variance-covariance for the parameters. The null was rejected for 10 out of the 40, indicating mild evidence of asymmetries in β . Finally, conditional asymmetries were added to the idiosyncratic variance

$$\sigma_{i,t}^2 = \alpha_{i,0} + \alpha_{i,1} RIV_{i,t-1} + \alpha_{i,2} \sigma_{i,t-1}^2 + \alpha_{i,3} RIV_{i,t-1} I(r_{f,t-1} < 0) + \alpha_{i,4} RIV_{i,t-1} I(\epsilon_{i,t-1} < 0)$$

where $\alpha_{i,3}$ allows for asymmetries due to the factor return and $\alpha_{i,4}$ allows for asymmetries due to the sign of the idiosyncratic shock. When tested with a Wald test, only 5 of the 40 series rejected the null $\alpha_{i,3} = \alpha_{i,4} = 0$, indicating that models for the idiosyncratic variance do not require conditional asymmetries. These results contrast sharply with what is typically found in standard low-frequency models, such as the conditional variance models in the cDCC. Low-frequency GJR-GARCH models,

$$\tilde{\sigma}_{i,t}^2 = \omega_i^D + \alpha_i^D r_{i,t-1}^2 + \beta_i^D \tilde{\sigma}_{i,t-1}^2 + \gamma_i^D r_{i,t-1}^2 I(r_{i,t-1} < 0)$$

were fit to the market return and the individual assets. In 35 of the 41 models fit, the estimate of γ_i^D was statistically significant. Moreover, in the market model, the estimate of α_f^D dropped from 0.08 to 0.00 ($\hat{\gamma}_f^D = 0.14$) when the asymmetric term was introduced. The median estimate of α_i^D was 0.09 when no asymmetry was included in the 40 individual firms; the introduction of asymmetries produced median estimates of 0.02 for α_i^D and 0.11 for γ_i^D . This difference between the low-frequency models and the high-frequency models is due to the decomposition which separates the market, which has a strong asymmetry from the idiosyncratic volatility, which does not. This additional parsimony is not applicable in standard low- (or high-) frequency cDCC-type models since the conditional variance estimated includes both components.

5 Conclusion

The paper introduces a new class of multivariate high-frequency based volatility models utilizing a factor structure for both the high- and low-frequency conditional covariances. This structure allows for parsimonious models to be fit while simplifying the task of incorporating high-frequency realized measures. The factor volatility, β and idiosyncratic volatility are modeled as separate HEAVY-type processes and the daily shocks in low-frequency models are replaced with the corresponding realized measures. This modification is supported by the empirical evidence that realized measures dominate daily measures for most equities. The Factor HEAVY is also validated as a covariance model using the specification tests. We show that this model improves on existing multivariate HEAVY models, and are especially useful for a large panels of assets.

In an empirical analysis of 40 systematically important U.S. financial firms, we found a one-factor model was adequate to capture the dynamics of the conditional covariance. The Factor HEAVY model performs better both in- and out-of-sample when compared to a leading low-frequency model. We find that factor HEAVY has a large advantage over cDCC in covariance matrix forecasting, and that the advantage appears in a variety of dimensions including the forecast copula, β and MES. The out-of-sample forecast ability is more pronounced at shorter forecast horizons which are the most relevant from a risk-management point-of-view. We consider extensions to two-factor models and check the robustness of the performance to some selected realized measures. We also examined a number of extensions to the base model including the role of asymmetries in the conditional variance, β and idiosyncratic variance, where we found that asymmetries are only needed for the conditional variance of the market.

References

- Acharya, V., Pedersen, L., Philippon, T. & Richardson, M. (2012), 'Measuring systemic risk'.
- Adrian, T. & Brunnermeier, M. (2009), 'Covar'. Federal Reserve Bank of New York Staff Reports 348.
- Aielli, G. P. (2013), 'Dynamic conditional correlations: on properties and estimation', *Journal of Business & Economic Statistics* **31**(3), 282–299.
- Aït-Sahalia, Y., Mykland, P. A. & Zhang, L. (2005), 'How often to sample a continuous-time process in the presence of market microstructure noise', *Review of Financial studies* **18**(2), 351–416.
- Andersen, T., Bollerslev, T., Diebold, F. & Labys, P. (2001), 'The distribution of realized exchange rate volatility', *Journal of the American Statistical Association* **96**, 42–55.
- Andersen, T., Bollerslev, T., Diebold, F. & Wu, J. (2006), 'Realized beta: Persistence and predictability'. in T. Fomby and D. Terrell, eds, *Advances in econometrics: Econometric analysis of economic and financial time series*, Vol. 20, Elsevier Science, pp. 1–39.
- Andersen, T. G., Bollerslev, T., Diebold, F. X. & Ebens, H. (2001), 'The distribution of realized stock return volatility', *Journal of Financial Economics* **61**(1), 43–76.
- Bali, T. & Engle, R. (2010), 'The intertemporal capital asset pricing model with dynamic conditional correlations', *Journal of Monetary Economics* **57**, 377–390.
- Bali, T., Engle, R. & Tang, Y. (2012), 'Dynamic conditional beta is alive and well in the cross-section of daily stock returns'. Working Paper.
- Bandi, F. & Russell, J. (2005), 'Realized covariation, realized beta and microstructure noise', *Unpublished paper, Graduate School of Business, University of Chicago*.
- Barndorff-Nielsen, O., Hansen, P., Lunde, A. & Shephard, N. (2008), 'Designing realized kernels to measure the ex post variation of equity prices in the presence of noise', *Econometrica* **76**, 1481–1536.
- Barndorff-Nielsen, O., Hansen, P., Lunde, A. & Shephard, N. (2009), 'Realized kernels in practice: trades and quotes', *Econometrics Journal* **12**, C1–C32.
- Barndorff-Nielsen, O., Hansen, P., Lunde, A. & Shephard, N. (2011), 'Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading', *Journal of Econometrics* **162**, 149–169.
- Barndorff-Nielsen, O. & Shephard, N. (2002), 'Econometric analysis of realized volatility and its use in estimating stochastic volatility models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**, 253–280.
- Barndorff-Nielsen, O. & Shephard, N. (2004), 'Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics', *Econometrica* **72**, 885–925.
- Bauwens, L., Laurent, S. & Rombouts, J. (2006), 'Multivariate garch models: a survey', *Journal of Applied Econometrics* **21**, 79–109.

- Bisias, D., Flood, M., Lo, A. & Valavanis, S. (2012), 'A survey of systemic risk analytics'. Office of Financial Research.
- Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics* **31**, 307–327.
- Bollerslev, T. & Wooldridge, J. (1992), 'Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances', *Econometric reviews* **11**, 143–172.
- Bollerslev, T. & Zhang, B. (2003), 'Measuring and modeling systematic risk in factor pricing models using high-frequency data', *Journal of Empirical Finance* **10**, 533–558.
- Braun, P. A., Nelson, D. B. & Sunier, A. M. (1995a), 'Good news, bad news, volatility, and betas', *Journal of Finance* **50**(5), 1575–1603.
- Braun, P., Nelson, D. & Sunier, A. (1995b), 'Good news, bad news, volatility, and betas', *Journal of Finance* **50**, 1575–1603.
- Brownlees, C. & Engle, R. (2012), Volatility, correlation and tails for systemic risk measurement. NYU Stern School of Business Working Paper.
- Christensen, K., Kinnebrock, S. & Podolskij, M. (2010), 'Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data', *Journal of Econometrics* **159**(1), 116–133.
- Diebold, F. & Mariano, R. (1995), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics* **13**, 253–263.
- Diebold, F. X. & Nerlove, M. (1989), 'The dynamics of exchange rate volatility: a multivariate latent factor arch model', *Journal of Applied Econometrics* **4**(1), 1–21.
- Engle, R. F. (2002), 'Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models', *Journal of Business & Economic Statistics* **20**, 339–350.
- Engle, R. & Kelly, B. (2012), 'Dynamic equicorrelation', *Journal of Business & Economic Statistics* **30**, 212–228.
- Engle, R., Ng, V. & Rothschild, M. (1990), 'Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills', *Journal of Econometrics* **45**, 213–237.
- Engle, R., Shephard, N. & Sheppard, K. (2008), 'Fitting vast dimensional time-varying covariance models', (403). NYU Working Paper.
- Ferson, W. & Harvey, C. (1993), 'The risk and predictability of international equity returns', *Review of Financial Studies* **6**, 527–566.
- Glosten, L. R., Jagannathan, R. & Runkle, D. E. (1993), 'On the relation between the expected value and the volatility of the nominal excess return on stocks', *The journal of finance* **48**(5), 1779–1801.
- Golosnoy, V., Gribisch, B. & Liesenfeld, R. (2012), 'The conditional autoregressive wishart model for multivariate stock market volatility', *Journal of Econometrics* **167**(1), 211–223.
- Gupta, A. K. & Nagar, D. K. (2000), *Matrix variate distributions*, Vol. 104, Chapman & Hall/CRC.

- Halbleib, R. & Voev, V. (2011), Forecasting covariance matrices: A mixed frequency approach, Technical report, Working Papers ECARES 2010-002, Universite Libre de Bruxelles.
- Hansen, P., Lunde, A. & Voev, V. (2012), 'Realized beta garch: A multivariate garch model with realized measures of volatility and covolatility'. CREATES Research Paper.
- Hansen, P. R., Lunde, A. & Voev, V. (2014), 'Realized Beta Garch: A Multivariate Garch Model With Realized Measures Of Volatility', *Journal of Applied Econometrics* pp. n/a–n/a.
URL: <http://dx.doi.org/10.1002/jae.2389>
- Jacod, J., Li, Y., Mykland, P., Podolskij, M. & Vetter, M. (2009), 'Microstructure noise in the continuous case: the pre-averaging approach', *Stochastic Processes and their Applications* **119**, 2249–2276.
- Jagannathan, R. & Wang, Z. (1996), 'The conditional capm and the cross-section of expected returns', *The Journal of Finance* **51**(1), 3–53.
- Laurent, S., Rombouts, J. V. & Violante, F. (2009), *Consistent ranking of multivariate volatility models*, CORE.
- Lewellen, J. & Nagel, S. (2006), 'The conditional capm does not explain asset-pricing anomalies', *Journal of Financial Economics* **82**(2), 289–314.
- Maheu, J. M. & Jin, X. (2013), 'Modelling realized covariances and returns', *Journal Of Financial Econometrics* **11**(2), 335–369.
- Newey, W. K. & McFadden, D. L. (1994), 'Large sample estimation and hypothesis testing'. In Robert F. Engle and Daniel L. McFadden(Eds.), *Handbook of Econometrics*, Volume 4, 2111–2245. North Holland.
- Ng, V., Engle, R. F. & Rothschild, M. (1992), 'A multi-dynamic-factor model for stock returns', *Journal of Econometrics* **52**(1), 245–266.
- Noureldin, D., Shephard, N. & Sheppard, K. (2012), 'Multivariate high-frequency-based volatility (heavy) models', *Journal of Applied Econometrics* **27**(6), 907–933.
- Patton, A. (2011), 'Volatility forecast comparison using imperfect volatility proxies', *Journal of Econometrics* **160**(1), 246–256.
- Patton, A. & Sheppard, K. (2009), 'Evaluating volatility and correlation forecasts'. in T.G. Andersen, R.A. Davis, J.P. Kreib and T. Mikosch, eds, *Handbook of Financial Time Series*, Springer, pp. 801–838.
- Patton, A. & Verardo, M. (2012), 'Does beta move with news? firm-specific information flows and learning about profitability', *Review of financial studies* **25**, 2789–2839.
- Shephard, N. & Sheppard, K. (2010), 'Realising the future: forecasting with high-frequency-based volatility (heavy) models', *Journal of Applied Econometrics* **25**, 197–231.
- Shephard, N. & Xiu, D. (2013), 'Econometric analysis of multivariate realised qml: estimation of the co-variation of equity prices under asynchronous trading', *Unpublished manuscript, University of Oxford*.
- White, H. (1996), *Estimation, inference and specification analysis*, number 22, Cambridge university press.

- Wooldridge, J. (1990), 'A unified approach to robust, regression-based specification tests', *Econometric Theory* **6**, 17–43.
- Zakoian, J.-M. (1994), 'Threshold heteroskedastic models', *Journal of Economic Dynamics and control* **18**(5), 931–955.
- Zhang, L. (2006), 'Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach', *Bernoulli* **12**, 1019–1043.

Appendices

A The Sample of Financial Firms

AET	Aetna	HRB	H&R Block
AFL	Aflac	JNS	Janus Capital
AIG	American International Group	JPM	JP Morgan Chase
ALL	Allstate Corp	KEY	Keycorp
AXP	American Express	LEH	Lehman Brothers
BAC	Bank of America	LM	Legg Mason
BBT	BB&T	MER	Merrill Lynch
BEN	Franklin Resources	MET	Metlife
BK	Bank of New York Mellon	MMC	Marsh & McLennan
C	Citigroup	MS	Morgan Stanley
CB	Chubb Corp	PGR	Progressive
CMA	Comerica inc	PNC	PNC Financial Services
COF	Capital One Financial	PRU	Prudential Financial
ETFC	E-Trade Financial	SCHW	Schwab Charles
EV	Eaton Vance	SLM	SLM Corp
FITB	Fifth Third Bancorp	STI	Suntrust Banks
FNM	Fannie Mae	STT	State Street
FRE	Freddie Mac	TROW	T. Rowe Price
GS	Goldman Sachs	UNH	Unitedhealth Group
HIG	Hartford Financial Group	USB	US Bancorp

B Proofs

B.1 Proof of Proposition 1

A $k \times k$ random symmetric positive definite matrix S is said to follow a Wishart distribution $W_k(\nu, \Sigma)$ with the degree of freedom ν and scale matrix Σ if the density is given by

$$\frac{|S|^{\frac{\nu-k-1}{2}}}{2^{\frac{\nu k}{2}} \Gamma_k(\frac{\nu}{2}) |\Sigma|^{\frac{\nu}{2}}} \exp\left(-\frac{\text{tr}(\Sigma^{-1}S)}{2}\right)$$

where $\Gamma_k(\frac{\nu}{2}) = \pi^{k(k-1)/4} \prod_{i=1}^k \Gamma(\frac{\nu+1-i}{2})$ is the multivariate gamma function.

As we assume $RM_t | \mathcal{F}_{t-1}^{\text{HF}} \sim W_{N+1}(\nu, M_t/\nu)$, the 2×2 realized covariance matrix

$$RM_{i,t} | \mathcal{F}_{t-1}^{\text{HF}} = \begin{bmatrix} RM_{f,t} & RM_{fi,t} \\ RM_{if,t} & RM_{ii,t} \end{bmatrix} | \mathcal{F}_{t-1}^{\text{HF}} \sim W_2(\nu, \frac{M_{i,t}}{\nu})$$

where $M_{i,t} = \begin{bmatrix} \mu_{f,t} & \lambda_{i,t}\mu_{f,t} \\ \lambda_{i,t}\mu_{f,t} & \lambda_{i,t}^2\mu_{f,t} + \mu_{i,t} \end{bmatrix}$. It follows that (Gupta & Nagar 2000)

$$R\beta_{i,t} | \mathcal{F}_{t-1}^{\text{HF}} = \frac{RM_{fi,t}}{RM_{f,t}} | \mathcal{F}_{t-1}^{\text{HF}} \sim T_{1,1}(\nu, \lambda_{i,t}, \mu_{f,t}^{-1}, \mu_{i,t})$$

$$RIV_{i,t} | \mathcal{F}_{t-1}^{\text{HF}} = (RM_{ii,t} - \frac{RM_{fi,t}^2}{RM_{f,t}}) | \mathcal{F}_{t-1}^{\text{HF}} \sim W_1(\nu - 1, \frac{\mu_{i,t}}{\nu})$$

where $T \sim T_{p,m}(\nu, M, \Sigma, \Omega)$ is the matrix variate t distribution with the density written as

$$\frac{\Gamma_p(\frac{\nu+m+p-1}{2})}{\pi^{\frac{mp}{2}} \Gamma_p(\frac{\nu+p-1}{2})} |\Sigma|^{-\frac{m}{2}} |\Omega|^{-\frac{p}{2}} |I_p + \Sigma^{-1}(T - M)\Omega^{-1}(T - M)'|^{-\frac{\nu+m+p-1}{2}}$$

Since $E[S] = \nu\Sigma$ and $E[T] = M$ if $S \sim W_k(\nu, \Sigma)$ and $T \sim T_{p,m}(\nu, M, \Sigma, \Omega)$,

$$E_{t-1} [R\beta_{i,t}] = \lambda_{i,t}$$

$$E_{t-1} [RIV_{i,t}] = (\nu-1/\nu)\mu_{i,t}$$

B.2 Proof of Proposition 3

The forecasts of $\sigma_{f,t+s}^2$, $\beta_{i,t+s}$ and $\sigma_{i,t+s}^2$ are derived by the same procedure, so we would just present for $\sigma_{f,t+s}^2$ for simplicity. By (5), we have

$$E_t [\sigma_{f,t+s}^2] = \theta_0 + \theta_1 E_t [RM_{f,t+s-1}] + \theta_2 E_t [\sigma_{f,t+s-1}^2] \quad (29)$$

The recursive equation (6) of the conditional mean of realized factor volatility gives rise to

$$E_t [\mu_{f,t+s}] = \theta_0^M + \theta_1^M E_t [RM_{f,t+s-1}] + \theta_2^M E_t [\mu_{f,t+s-1}] \quad (30)$$

Since $E_t(\mu_{f,t+s}) = E_t(E_{t+s-1}(RM_{f,t+s})) = E_t(RM_{f,t+s})$ by the law of iterated expectations, (30) can be rewritten as

$$E_t [RM_{f,t+s}] = \theta_0^M + (\theta_1^M + \theta_2^M) E_t [RM_{f,t+s-1}]$$

when $s \geq 2$. Thus,

$$E_t [RM_{f,t+s}] = \theta_0^M \frac{1 - (\theta_1^M + \theta_2^M)^{s-1}}{1 - (\theta_1^M + \theta_2^M)} + (\theta_1^M + \theta_2^M)^{s-1} \mu_{f,t+1} \quad (31)$$

By Substituting (31) for $E_t [RM_{f,t+s-1}]$ into (29) and forward iterating, we finally have the close-form of s -step forecast of (14).

B.3 Proof of Proposition 4

Here we only present the proof of (19), since (21) can be similarly proved. We partition Σ_t as follows

$$\begin{aligned}\Sigma_t &= \begin{bmatrix} \sigma_{f,t}^2 & \beta_{1,t}\sigma_{f,t}^2 & \beta_{2,t}\sigma_{f,t}^2 & \cdots & \beta_{N,t}\sigma_{f,t}^2 \\ \beta_{1,t}\sigma_{f,t}^2 & \beta_{1,t}^2\sigma_{f,t}^2 + \sigma_{1,t}^2 & \beta_{1,t}\beta_{2,t}\sigma_{f,t}^2 & \cdots & \beta_{1,t}\beta_{N,t}\sigma_{f,t}^2 \\ \beta_{2,t}\sigma_{f,t}^2 & \beta_{2,t}\beta_{1,t}\sigma_{f,t}^2 & \beta_{2,t}^2\sigma_{f,t}^2 + \sigma_{2,t}^2 & \cdots & \beta_{2,t}\beta_{N,t}\sigma_{f,t}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \beta_{N,t}\sigma_{f,t}^2 & \beta_{N,t}\beta_{1,t}\sigma_{f,t}^2 & \beta_{N,t}\beta_{2,t}\sigma_{f,t}^2 & \cdots & \beta_{N,t}^2\sigma_{f,t}^2 + \sigma_{N,t}^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_{f,t}^2 & \sigma_{f,t}^2\beta_t' \\ \sigma_{f,t}^2\beta_t & \sigma_{f,t}^2\beta_t\beta_t' + \sigma_t^2 \end{bmatrix}\end{aligned}\quad (32)$$

where $\beta_t = (\beta_{1,t}, \dots, \beta_{N,t})'$ and

$$\sigma_t^2 = \begin{bmatrix} \sigma_{1,t}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{2,t}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{N,t}^2 \end{bmatrix}.$$

Using the partitioned inverse theorem,

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}B(D - CA^{-1}B)^{-1}CA^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -(D - CA^{-1}B)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix},$$

the inverse of (32) can be written as

$$\Sigma_t^{-1} = \begin{bmatrix} \sigma_{f,t}^2 & \sigma_{f,t}^2\beta_t' \\ \sigma_{f,t}^2\beta_t & \sigma_{f,t}^2\beta_t\beta_t' + \sigma_t^2 \end{bmatrix}^{-1} = \begin{bmatrix} (\sigma_{f,t}^2)^{-1} + \beta_t'(\sigma_t^2)^{-1}\beta_t & -\beta_t'(\sigma_t^2)^{-1} \\ -(\sigma_t^2)^{-1}\beta_t & (\sigma_t^2)^{-1} \end{bmatrix}\quad (33)$$

Define $\tilde{r}_t = (r_{1,t}, \dots, r_{N,t})$ as the returns on the N assets, then

$$\begin{aligned}& [r_{f,t} \ r_{1,t} \ \dots \ r_{N,t}] \Sigma_t^{-1} [r_{f,t} \ r_{1,t} \ \dots \ r_{N,t}]' \\ &= [r_{f,t} \ \tilde{r}_t] \begin{bmatrix} (\sigma_{f,t}^2)^{-1} + \beta_t'(\sigma_t^2)^{-1}\beta_t & -\beta_t'(\sigma_t^2)^{-1} \\ -(\sigma_t^2)^{-1}\beta_t & (\sigma_t^2)^{-1} \end{bmatrix} \begin{bmatrix} r_{f,t} \\ \tilde{r}_t' \end{bmatrix} \\ &= \left((\sigma_{f,t}^2)^{-1} + \beta_t'(\sigma_t^2)^{-1}\beta_t \right) r_{f,t}^2 - 2r_{f,t}\beta_t'(\sigma_t^2)^{-1}\tilde{r}_t' + \tilde{r}_t'(\sigma_t^2)^{-1}\tilde{r}_t \\ &= \frac{r_{f,t}^2}{\sigma_{f,t}^2} + \sum_{i=1}^N \frac{(r_{i,t} - \beta_{i,t}r_{f,t})^2}{\sigma_{i,t}^2}\end{aligned}\quad (34)$$

The determinant of the invertible partitioned matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ is $|A||D - CA^{-1}B|$, and so

$$|\Sigma_t| = \sigma_{f,t}^2 \prod_{i=1}^N \sigma_{i,t}^2\quad (35)$$

Substituting (34) and (35) into (18), one can finally get (19).